

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

---

SCUOLA DI SCIENZE  
Corso di Laurea in Informatica

**Analisi delle funzionalita' respiratorie**  
**Monitoraggio della respirazione attraverso uno stetoscopio**  
**elettronico**

**Tesi di Laurea in Architettura degli Elaboratori**

**Relatore:**  
**Chiar.mo Prof.**  
**Vittorio Ghini**

**Presentata da:**  
**Laura Baldini**

**Sessione III**  
**Anno Accademico 2011/2012**



# Indice

<b>Elenco delle figure</b>	<b>7</b>
<b>Introduzione</b>	<b>9</b>
0.1    Contenuto della tesi . . . . .	10
<b>I    Prerequisiti</b>	<b>11</b>
<b>1    Respirazione e apparato respiratorio</b>	<b>13</b>
1.1    Anatomia dell'apparato respiratorio . . . . .	13
1.2    Respirazione . . . . .	14
1.2.1    Meccanica della respirazione . . . . .	14
1.3    Diagnosi delle malattie dell'apparato respiratorio . . . . .	15
1.3.1    Spirometro con pneumotacografo . . . . .	15
1.3.2    Polisonnografia . . . . .	15
1.3.3    Auscultazione . . . . .	15
1.3.4    Stetoscopio . . . . .	15
1.3.5    Invasività clinica . . . . .	16
1.4    Patologia dell'apparato respiratorio . . . . .	16
1.4.1    Apnee del sonno . . . . .	16
1.4.2    Arresto respiratorio . . . . .	18
1.5    Base funzionale dei suoni respiratori . . . . .	18
1.6    Analisi acustica dei suoni respiratori . . . . .	18
1.7    Schemi di respirazione . . . . .	20
<b>2    Acustica e audio digitale</b>	<b>23</b>
2.1    L'acustica e il suono . . . . .	23
2.2    Onde sonore . . . . .	23
2.3    Suono analogico e digitale . . . . .	24
2.3.1    Discretizzazione . . . . .	25

2.4	segnale . . . . .	27
2.5	Analisi nel dominio del tempo e nel dominio della frequenza . . . . .	28
2.6	Analisi armonica e trasformata di Fourier . . . . .	29
2.6.1	Serie di Fourier . . . . .	30
2.6.2	Trasformata di Fourier . . . . .	30
2.6.3	Trasformata discreta di Fourier . . . . .	31
2.7	Wavelet . . . . .	31
2.7.1	Trasformata wavelet . . . . .	32
<b>3</b>	<b>Filtri digitali</b>	<b>35</b>
3.1	Denoising . . . . .	35
3.2	Classificazione dei filtri digitali . . . . .	35
3.2.1	Categorie di filtri . . . . .	36
3.3	Filtrare i suoni respiratori . . . . .	37
<b>4</b>	<b>Metodi di base</b>	<b>41</b>
4.1	Beat detection e algoritmi di onset detection . . . . .	41
4.1.1	Preprocessing . . . . .	42
4.1.2	Reduction . . . . .	42
4.1.3	Postprocessing . . . . .	43
4.1.4	Thresholding . . . . .	43
4.1.5	Peak picking . . . . .	44
4.2	Clustering . . . . .	44
4.2.1	Algoritmi di clustering . . . . .	44
4.3	Reti neurali . . . . .	46
4.4	Dimensione frattale . . . . .	47
4.5	Catene di triple Markoviane . . . . .	47
4.5.1	Nozioni di base di probabilità . . . . .	47
4.5.2	Processo di Markov . . . . .	49
4.5.3	Catena di Markov . . . . .	49
4.5.4	Modelli nascosti di Markov . . . . .	49
4.5.5	Catena di coppie markoviane . . . . .	50
4.5.6	Catena di triple markoviane . . . . .	50
<b>II</b>	<b>Stato dell'arte e analisi delle metodologie</b>	<b>51</b>
<b>5</b>	<b>Stato dell'arte</b>	<b>53</b>
5.1	Algoritmi di riconoscimento delle fasi e dei cicli respiratori . . . . .	53

5.1.1	Automated respiratory phase and onset detection using only chest sound signal . . . . .	54
5.1.2	Breath Analysis of Respiratory Flow using Tracheal Sounds . . .	55
5.1.3	Acoustical respiratory signal analysis and phase detection . . . .	57
5.1.4	Computerized acoustical respiratory phase detection without air-flow measurement . . . . .	58
5.1.5	Respiratory onset detection using variance fractal dimension . . .	59
5.1.6	Automated respiratory phase detection by acoustical means . . .	60
5.1.7	A Software Toolkit for Acoustic Respiratory Analysis . . . . .	62
5.2	Classificazione degli algoritmi . . . . .	65
5.3	Conclusioni . . . . .	66
<b>6</b>	<b>Analisi delle Metodologie</b>	<b>67</b>
6.1	Metodo elementare . . . . .	68
6.2	Beat detection . . . . .	68
6.2.1	Scelta dell'algoritmo . . . . .	68
6.3	Pattern recognition e apprendimento automatico . . . . .	69
6.3.1	Reti neurali . . . . .	69
6.4	Clustering . . . . .	70
<b>7</b>	<b>Meccanismi di valutazione</b>	<b>73</b>
7.1	Classificazione dello spazio dell'input . . . . .	73
7.2	Test valutativi . . . . .	74
<b>III</b>	<b>Conclusioni</b>	<b>75</b>
	<b>Conclusioni</b>	<b>77</b>
7.3	Sviluppi futuri . . . . .	78
<b>8</b>	<b>Bibliografia</b>	<b>81</b>
8.1	Sitografia . . . . .	84



# Elenco delle figure

1.1	Apparato respiratorio . . . . .	14
1.2	Pattern respiratori anormali . . . . .	21
2.1	Onda sonora . . . . .	24
2.2	Discretizzazione . . . . .	25
2.3	Campionamento . . . . .	26
2.4	Quantizzazione . . . . .	27
2.5	Envelope di un segnale . . . . .	28
2.6	Grafici dei domini tempo/frequenza . . . . .	29
2.7	Scomposizione di un segnale . . . . .	30
2.8	Onda wavelet . . . . .	32
2.9	Applicazioni della wavelet al segnale . . . . .	33
3.1	Filtro passa alto . . . . .	36
3.2	Filtro passa basso . . . . .	36
3.3	Filtro passa banda . . . . .	37
3.4	Filtro soppressore di banda . . . . .	37
3.5	Funzione finestra . . . . .	39
4.1	Algoritmo di beat detection . . . . .	42
5.1	Suono respiratorio dopo l'uso di filtri . . . . .	55
6.1	Diagramma di implementazione . . . . .	67
6.2	Forma d'onda del suono respiratorio . . . . .	71
7.1	Sorgente dei suoni rilevati da uno stetoscopio . . . . .	73





# Introduzione

Questa tesi si concentra sull'analisi delle funzionalità respiratorie attraverso lo studio del suono per poter affrontare un diffuso problema dell'apparato respiratorio : la sindrome dell'apnea notturna o apnea del sonno. Questa sindrome consiste nell'interruzione dell'attività respiratoria durante il sonno, queste interruzioni vengono definite come eventi di apnea con una durata di diversi secondi, in caso di una sindrome leggera possono verificarsi dai cinque ai quindici eventi in un'ora, mentre per una forma più grave si può arrivare fino a trenta eventi.

L'apnea notturna, se non viene curata, aumenta il rischio di problemi cardiaci, infarti ed eventi cerebrovascolari. Lo studio [3] conclude che c'è una forte associazione tra la sindrome di apnea del sonno e ictus. Inoltre tale patologia contribuisce a creare un senso perenne di sonnolenza e spossatezza. Si stima che più della metà delle persone affette da tale patologia non ne siano al corrente [1], inoltre è utile considerare che la diagnosi di questa patologia necessita di strumenti costosi o scomodi che richiedono al paziente di trascorrere la notte in un centro specializzato o comunque in una struttura contenente un'apparecchiatura adeguata e purtroppo queste strutture non sono molto diffuse.

Alcuni studi analizzando i dati di alcuni soggetti deceduti a causa di eventi cardiovascolari acuti affetti da forme medie o gravi di sindrome da apnea del sonno hanno concluso che la maggior parte di tali soggetti è morta durante il sonno.

In questa tesi, dopo una accurata analisi dell'ambiente, si suggerisce un metodo di monitoraggio del respiro e di diagnosi di sindrome da apnea del sonno: attraverso uno stetoscopio elettronico posizionato sul paziente durante le ore notturne il sistema registra la respirazione del soggetto e deve vigilare su esso individuando gli eventi di apnea. Il sistema deve svegliare il paziente tramite dei semplici allarmi sonori nel caso in cui si presenti un'apnea che possa compromettere il soggetto, così da indurlo a svegliarsi per evitare gravi conseguenze.

## 0.1 Contenuto della tesi

La tesi si divide in tre parti:

1. La prima contiene i prerequisiti necessari alla comprensione delle parti successive. In questa parte abbiamo deciso di parlare di tutti gli argomenti necessari, con un livello di sintesi proporzionale all'importanza dell'argomento. Il livello di sintesi dovrebbe essere proporzionale a quanto ci aspettiamo che il lettore sappia. Tuttavia non è possibile prevedere bene quali siano i nostri lettori e quali conoscenze abbiano.
2. La seconda parte contiene: una presentazione del problema; un riassunto dello stato dell'arte; una analisi e una discussione delle metodologie in atto e in potenza per la risoluzione del problema.
3. L'ultima parte invece contiene le conclusioni effettuate dopo l'analisi del problema.

Di seguito entriamo più nel dettaglio dei singoli capitoli.

Il capitolo 1 dà una sintesi delle nozioni mediche necessarie alla comprensione del resto della tesi. In particolare parla brevemente: dell'anatomia dell'apparato respiratorio; di cosa sono i suoni respiratori; di come e dove questi ultimi hanno origine e che caratteristiche acustiche hanno; di alcune patologie del sistema respiratorio e di quali strumenti diagnostici disponiamo.

Il capitolo 2 si occupa di audio digitale in generale.

Il capitolo 3 tratta i filtri digitali in generale e alcuni filtri digitali usati nello stato dell'arte con un orientamento soprattutto sull'utilità intuitiva del filtro piuttosto che sugli aspetti matematici e di implementazione.

Il capitolo 4 descrive le tecniche usate o usabili nella soluzione del problema cioè: beat detection, reti neurali, dimensione frattale, clustering, catene di Markov.

Il capitolo 5 descrive lo stato dell'arte nella soluzione del problema.

Il capitolo 6 descrive le metodologie usate e alcune delle metodologie usabili per la soluzione del problema.

Il capitolo 7 descrive i metodi valutativi per un buon sistema di risoluzione.

# **Parte I**

## **Prerequisiti**



# Capitolo 1

## Respirazione e apparato respiratorio

### 1.1 Anatomia dell'apparato respiratorio

L'*apparato respiratorio* (o anche sistema respiratorio) è un sistema biologico atto alla respirazione che permette gli scambi gassosi di ossigeno e anidride carbonica tra il sangue e l'ambiente esterno. Negli esseri umani l'anatomia funzionale dell'apparato respiratorio include [55]:

**Vie aeree** Le vie aeree sono cavità in cui le sostanze gassose, vengono trasportate da o verso i polmoni. Le seguenti parti del corpo sono vie aeree: naso esterno, cavità orale, faringe, laringe, trachea. La figura 1.1 è una rappresentazione più esaustiva dell'apparato respiratorio.

**Polmoni** I polmoni sono l'organo essenziale per la respirazione. La sua principale funzione è quella di trasportare l'ossigeno dell'ambiente circostante al sangue e di espellere l'anidride carbonica. I polmoni contengono delle piccole sacche d'aria chiamate alveoli. Attraverso i capillari degli alveoli avviene lo scambio per diffusione di ossigeno e anidride carbonica tra il sangue dell'organismo e l'aria contenuta negli alveoli. Il polmone destro è diviso in tre lobi: superiore, medio e inferiore; mentre quello sinistro è diviso in due lobi: uno superiore e uno inferiore.

**Muscoli respiratori** I muscoli respiratori sono il diaframma, i muscoli intercostali, i muscoli addominali, lo sternocleidomastoideo e i muscoli scaleni. Questi muscoli hanno il ruolo di causare l'espansione polmonare.

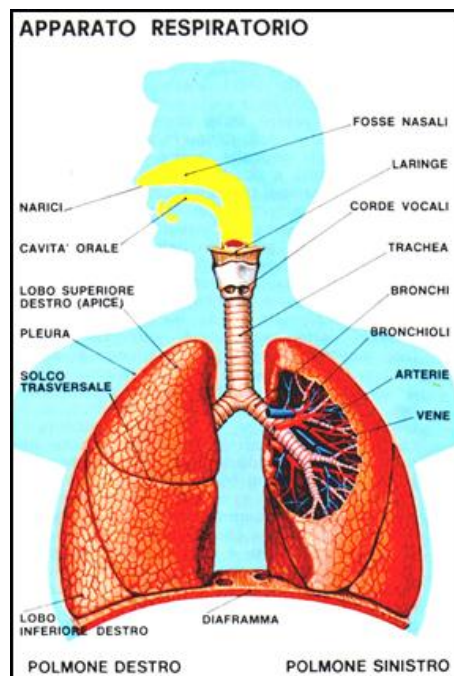


Figura 1.1: Apparato respiratorio umano [49]

## 1.2 Respirazione

La *respirazione* è il processo attraverso il quale l'organismo scambia aria tra i polmoni e l'ambiente circostante. La respirazione permette di acquisire ossigeno nel sangue che viene poi usato nel metabolismo e permette di eliminare i residui gassosi del metabolismo come l'anidride carbonica. Durante la respirazione si scambia anche vapore acqueo. Il termine usato per indicare il respiro in condizioni normali è *eupnea*.

### 1.2.1 Meccanica della respirazione

La respirazione è ciclica e un ciclo di respirazione ha atto in tre fasi [14]:

**Inspirazione** Durante questa fase l'aria viene introdotta nei polmoni. L'inspirazione avviene grazie alla contrazione dei muscoli intercostali e del diaframma, tale contrazione provoca un aumento di volume polmonare e una diminuzione della pressione intrapleurica: ne consegue un'aspirazione dell'aria nei polmoni.

**Espirazione** Durante questa fase l'aria viene espulsa dai polmoni. Il volume toracico diminuisce, i polmoni vengono compressi e l'aria espulsa. L'efficienza della espirazione è soprattutto passiva, a differenza dell'inspirazione, ed è dovuta al ritorno elastico del tessuto polmonare.

**Pausa** Tra una espirazione e l'inspirazione successiva ci può essere una pausa di durata variabile.

## 1.3 Diagnosi delle malattie dell'apparato respiratorio

### 1.3.1 Spirometro con pneumotacografo

Lo spirometro è uno strumento utilizzato per misurare i volumi d'aria polmonari. È composto da un sensore collegato a un boccaglio, attraverso il quale il paziente respira e da una strumentazione che misura i movimenti di aria provocati dal soggetto. Lo spirometro con pneumotacografo è un tipo di spirometro dotato di un setto poroso che causa una differenza di pressione. Questa differenza pressoria viene misurata da un manometro e poi riconvertita in un segnale proporzionale al flusso generato [61] [64].

### 1.3.2 Polisonnografia

La *polisonnografia* è una tecnica diagnostica che consiste nella registrazione simultanea di più parametri fisiologici durante la notte. Normalmente nel corso del test vengono registrati: il flusso d'aria della respirazione, il livello di ossigeno nel sangue, la posizione del corpo, attività cerebrale (attraverso un elettroencefalogramma), il movimento degli occhi (attraverso un elettrooculogramma) e l'attività cardiaca (attraverso un elettrocardiogramma) [53].

### 1.3.3 Auscultazione

L'*auscultazione* è un sistema di diagnosi che consiste nell'ascoltare i suoni interni del corpo. La parte di interesse si riferisce al termine con il significato di ascolto, attraverso uno stetoscopio, dei suoni prodotti dall'apparato respiratorio. Lo stetoscopio può essere posizionato in varie parti del torace, della schiena e del collo [26].

### 1.3.4 Stetoscopio

Lo stetoscopio è uno strumento medico utile all'auscultazione delle parti interne del corpo e soprattutto del torace. Ci sono due tipi principali di stetoscopio:

**Stetoscopio acustico** Lo stetoscopio acustico funziona tramite la trasmissione di suoni provenienti dal petto attraverso dei canali contenenti aria fino alle orecchie dell'uditore. Uno dei problemi degli stetoscopi acustici è il basso volume del suono trasmesso. Ne consegue, molto spesso, una difficoltà nell'eseguire una diagnosi precisa.

**Stetoscopio elettronico** L'innovazione tecnologica nel campo degli stetoscopi consente oggi la rilevazione di una più ampia gamma di suoni e una maggior qualità d'ascolto, con la possibilità di ottenere registrazioni e riproduzioni di grande fedeltà dei suoni. A differenza degli stetoscopi acustici, che sono tutti basati sullo stesso metodo di funzionamento, gli stetoscopi elettronici variano molto tra un modello e l'altro. I più semplici funzionano grazie a un microfono applicato al petto del paziente, ma di contro, risentono del suono dell'ambiente esterno che spesso causa interferenze. Alcuni stetoscopi elettronici forniscono direttamente un audio in output e si possono usare insieme a un dispositivo esterno di registrazione. Altri possono trasmettere il segnale via wireless o bluetooth. Un dispositivo (per esempio un computer) riceve questo segnale e può fare vari tipi di analisi. Altri modelli, più complessi, trasformano le onde sonore in impulsi elettrici, così da poter essere amplificate per un migliore ascolto. Recentemente, alcuni stetoscopi elettronici sono stati dotati di filtri, allo scopo di eliminare le interferenze sonore esterne e in più sono in grado di poter selezionare il range di frequenza in modo da poter ascoltare separatamente o meno i suoni cardiaci da quelli respiratori [59].

### 1.3.5 Invasività clinica

Un parametro molto importante per valutare un sistema diagnostico è l'invasività clinica, questa si riferisce alla possibilità che l'esame finisca per compromettere ulteriormente lo stato di salute del soggetto. Siamo in presenza di un esame invasivo per esempio, nel caso in cui l'esame possa portare agenti contaminanti (virus, batteri, tossine, sporcizia) all'interno del diretto interessato e quindi causare una infezione che aggravi le condizioni del paziente. L'invasività di un meccanismo diagnostico si nota anche nei casi in cui un piccolo errore nella procedura procuri danni al paziente [48].

Tra i sistemi diagnostici descritti in questa sezione il metodo meno invasivo è l'auscultazione.

## 1.4 Patologia dell'apparato respiratorio

### 1.4.1 Apnee del sonno

La *sindrome da apnea del sonno* è un disordine del sonno caratterizzato da ripetute apnee o ipopnee durante il sonno [3]. Una *apnea* è una pausa di durata anormale nella respirazione che supera i dieci secondi [17] e può durare anche alcuni minuti [27]. Una *ipopnea* è un evento caratterizzato da respirazione insufficiente, più precisamente una ipopnea si verifica quando il flusso d'aria si riduce di almeno il 30% per un minimo di



dieci secondi e la desaturazione di ossigeno nel sangue è di almeno il 4% [17]. Ci sono tre forme di apnea del sonno:

**Apnea centrale** La respirazione è interrotta per via di un mancato movimento dei muscoli respiratori, quindi il volume dei polmoni rimane invariato.

**Apnea ostruttiva** La respirazione è interrotta a causa di un blocco fisico nelle vie aeree nonostante persistano movimenti respiratori [5].

**Apnea mista** La respirazione è soggetta a entrambi i tipi di apnea appena descritti.

Le apnee del sonno si verificano sia nei bambini che negli adulti. I soggetti affetti da apnea del sonno possono manifestare i seguenti sintomi: eccessiva sonnolenza durante il giorno, tempi di reazione lenti, problemi alla vista, indebolimento delle funzioni del fegato e altro. Inoltre gravi forme di apnee del sonno ostruttive aumentano in modo significativo il rischio di eventi cardiovascolari fatali [17] [32] [25].

L'*indice di apnea-ipopnea (apnea-hypopnea index AHI)* è definito come il numero di eventi di apnea e di ipopnea rispetto al tempo di sonno. L'AHI è un indicatore della gravità della sindrome di apnea del sonno e i suoi valori sono categorizzati tipicamente in: leggera da 5 a 15 episodi all'ora; moderata da 15 a 30 episodi all'ora o severa oltre i 30 episodi all'ora.

Lo studio [3] conclude che c'è una forte associazione tra la sindrome da apnea del sonno e l'ictus. In particolare dimostra che un indice da moderato a severo di AHI è associato a un alto rischio di ictus e ipotizza che la sindrome da apnea del sonno contribuisca allo sviluppo di un ictus. Purtroppo resta ancora da scoprire se esiste un pattern respiratorio specifico che precede immediatamente un ictus.

Lo studio [12] prende in esame i polisonnogrammi e i certificati di morte di alcune persone che sono decedute a causa di una malattia cardiaca improvvisa. Le persone sono state divise in due gruppi: le persone del primo gruppo soffrivano di sindrome da apnea notturna mentre le persone del secondo no. Si è riscontrato che la maggior parte delle persone appartenenti al primo gruppo sono morte durante il sonno al contrario di quelle del secondo. Lo studio conclude che la gravità della sindrome da apnea del sonno è direttamente proporzionale al rischio di morte improvvisa per malattie cardiache durante il sonno.

Episodi acuti di apnea o ipopnea possono indurre: ipossiemia, aumento degli impulsi nel sistema nervoso simpatico, aumento brusco nella pressione sanguigna, aumento dello stress delle pareti cardiache, aritmie cardiache, ipercoagulabilità, stress ossidativo vascolare, infiammazioni sistemiche e altro. Questo potrebbe spiegare i dati osservati dallo studio. Resta un problema aperto quello di stabilire se nel primo gruppo di persone, la morte è immediatamente preceduta da un evento di apnea o ipopnea grave.

### 1.4.2 Arresto respiratorio

Un arresto respiratorio è l'interruzione della respirazione normale. È molto probabile che si verifichino traumi cerebrali se l'arresto respiratorio dura più di tre minuti e la morte è quasi certa se questo dura più di cinque minuti.

## 1.5 Base funzionale dei suoni respiratori

Secondo [29] i suoni normali che si possono sentire sul petto di un soggetto nella fase di inspirazione, vengono generati soprattutto nella parte lobare delle vie respiratorie. I suoni respiratori sono generati da turbolenze dell'aria nelle vie respiratorie. Le caratteristiche dei suoni sono molto variabili, si possono notare differenze da persona a persona che dipendono dal peso, dall'età, dallo stato di salute e altri fattori. Si nota anche una variabilità rispetto alla densità del gas respirato. Lo studio [19] nota che i suoni inspiratori vengono prodotti in modo predominante nella zone periferiche dei polmoni, mentre i suoni espiratori vengono prodotti nelle zone più centrali.

## 1.6 Analisi acustica dei suoni respiratori

Lo studio [13] analizza l'energia spettrale dei suoni respiratori normali, ascoltati sul petto di un soggetto. Conclude che questi seguono uno schema caratteristico, in particolare che la potenza decresce in modo esponenziale con la frequenza nell'intervallo dai 75 ai 2000Hz, dopo quest'ultimo valore la potenza è trascurabile. Secondo [9] il suono registrato sul petto di soggetti sani ha una ampiezza massima (in Decibel) attorno ai 250Hz e decresce rapidamente fino ad arrivare a un livello di ampiezza trascurabile attorno a una frequenza di circa 1000Hz. Nel caso dei soggetti malati invece i picchi di frequenze sono più alti.

[37] [18] [16] classificano i suoni respiratori nel modo seguente (riassunto nella tabella 1.1):

**Suoni respiratori normali** I suoni respiratori normali o fisiologici sono i suoni prodotti dal sistema respiratorio di un soggetto sano. Questi suoni sono più intensi in fase inspiratoria e decrescono in fase espiratoria (nonostante quest'ultima sia più lunga di quella inspiratoria) e ciò è dovuto al fatto che l'inspirazione avviene mediante contrazione muscolare, generando un flusso d'aria più veloce, rispetto all'espirazione, che è un fenomeno passivo. Si dividono in:

**Murmori vescicolari** Sono più intensi in fase inspiratoria e decrescono in fase espiratoria. Questi suoni si generano per l'azione di filtro degli alveoli sull'aria in arrivo.

**Suoni broncovescicolari** Si generano nelle zone di passaggio dall'albero bronchiale agli alveoli, quindi per mescolanza di rumore bronchiale e alveolare.

**Suono o soffio bronchiale** Ha una fase inspiratoria breve e una fase espiratoria lunga, separati da una pausa. Corrisponde alla proiezione dei bronchi sul torace e si sente particolarmente bene nella parte anteriore del paziente.

I suoni normali hanno una frequenza nell'intervallo dai 100 ai 1000Hz. A causa dei suoni prodotti dal movimento dei muscoli del torace e dal diaframma e a causa di quelli prodotti dall'apparato cardiocircolatorio, i suoni dall'apparato respiratorio di solito non vengono studiati a frequenze minori di 60Hz. I suoni respiratori dipendono molto dal flusso d'aria quindi lo spettro dei suoni dovrebbe essere riportato in funzione di esso. Inoltre è importante determinare lo spettro di frequenza corrispondente a un flusso d'aria pari a zero, con lo scopo di determinare il rumore di fondo [8].

**Suoni anormali** I suoni respiratori anormali o patologici, sono suoni accidentali che non fanno parte del normale ciclo di respirazione. Ci sono due tipi di suoni anormali:

**Continui** Tra i suoni anormali continui si distinguono:

**Ronchi o gemiti** Questi suoni sono spesso gravi e sembrano dei rantoli, la loro frequenza dominante è 200Hz, hanno quindi una bassa tonalità. Sono segno di broncocostrizione e si generano per il passaggio dell'aria in vie aeree ristrette per la presenza di muco o broncospasmo.

**Rantolo secco (wheezes)** Questi suoni hanno una frequenza dominante attorno ai 100Hz e hanno una durata maggiore di 100ms. Questi suoni sono un segnale caratteristico di una malattia ostruttiva polmonare [8].

**Sibili** Questi suoni sono dei *wheezes* molto più forti e sono la conseguenza di una ostruzione dinamica nella laringe o nella trachea. L'energia di questi suoni è concentrata in massima parte su una frequenza che si aggira attorno ai 1000Hz.

**Discontinui** Tra i suoni anormali discontinui si notano:

**Crackles** questi suoni hanno una durata minore di 10ms e cadono in un vasto spettro di frequenze, tra 200 e 2000Hz. Di solito sono indice di malattie cardiorespiratorie.

**Fine crackles** Sono i rantoli fini, anche detti rantoli crepitanti e hanno una tonalità alta. Sono dovuti alla ritardata apertura degli alveoli. Talvolta a essi si sovrappongono gli sfregamenti pleurici, il cui suono respiratorio prodotto è molto simile.

**Coarse crackles** Sono i rantoli subcrepitanti, di una tonalità bassa, il cui respiro prodotto si modifica con la tosse. Sono segno di bronchite e si generano per il passaggio dell'aria attraverso il muco.

**Squawks** questi suoni sono dei *wheezes* di breve durata.

Suono	Durata	Spettro di frequenza	Normalità	Continuità
suoni normali bronchiali	-	100 – 1000Hz	normale	continuo
suoni normali vescicolari	-	100 – 1000Hz	normale	continuo
ronchi	-	~ 200Hz	anormale	continuo
wheeze	100 – 250ms	~ 100Hz	anormale	continuo
stridor	-	~ 1000Hz	anormale	continuo
crackles	-	-	anormale	discontinuo
squawks	-	-	anormale	discontinuo

Tabella 1.1: Classificazione dei suoni respiratori

I risultati ottenuti da [19] dicono che in media, a parità di flusso i suoni inspiratori hanno una intensità di circa 10db maggiore rispetto ai suoni espiratori. Inoltre i suoni registrati da microfoni in prossimità della trachea sono più intensi di quelli registrati da microfoni sul petto.

I rantoli di solito vengono registrati in prossimità della trachea, dove gli effetti dell'attenuazione di tali suoni da parte del corpo sono minimi. D'altro canto, i suoni discontinui accidentali generati a causa di patologie polmonari vengono riconosciuti meglio sulla parte bassa dei polmoni [33].

**Rumore** Ci sono altri suoni fisiologici che complicano l'auscultazione dei suoni respiratori: i suoni cardiovascolari e i suoni gastrointestinali. Nel caso dell'auscultazione dei suoni tracheali abbiamo un'altra importante fonte di disturbo e cioè i suoni dovuti alla deglutizione della saliva.

## 1.7 Schemi di respirazione

Lo studio [35] analizza gli schemi o pattern di respirazione anormali. I soggetti coinvolti nello studio sono 65 e hanno età dai 20 agli 81 anni. Vengono presi in considerazione

vari parametri tra i quali: la frequenza di respirazione e il tempo di inspirazione. Lo studio conclude che la media delle frequenze di respirazione è  $16.6 \pm 2.8$  respiri al minuto mentre la durata media dei tempi di inspirazione è di  $1.62 \pm 0.31$  secondi.

Esistono vari schemi di respirazione anormale, più in dettaglio, analizzando la figura 1.2 dall'alto verso il basso si ha:

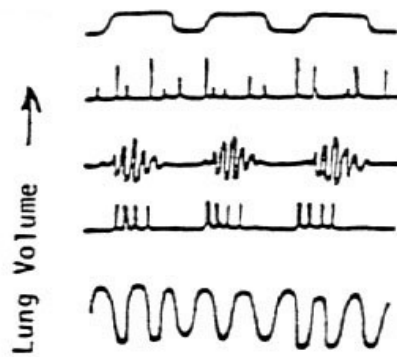


Figura 1.2: Pattern respiratori anormali [51]

**Apneosi** L'apneosi è uno schema di respirazione caratterizzato da profondi annaspamenti durante l'inspirazione, dopo una inspirazione avviene una lunga pausa seguita da un rilascio insufficiente di aria.

**Respiro agonico** Il respiro agonico è uno schema anormale di respirazione caratterizzato da boccheggiamento e da una riduzione estrema della frequenza degli atti respiratori fino al loro totale arresto. L'uso corretto del termine si deve restringere all'ultimo respiro prima della morte.

**Respiro di Biot** Il respiro di Biot (chiamato anche respiro intermittente o respirazione atassica) è uno schema di respirazione caratterizzato da periodi di respirazione brevi e superficiali seguiti da fasi di apnea di durata variabile, ma in genere tra i 10 ed i 30 secondi.

**Respiro di Cheyne Stokes** Il respiro di Cheyne Stokes è uno schema di respirazione anormale nel quale in un primo momento il respiro è frequente e profondo, successivamente la frequenza respiratoria modifica diminuendo gradualmente fino a portare a una apnea. Lo schema si ripete e ogni ciclo di solito dura dai 30 secondi ai 2 minuti.

**Respiro di Kussmaul** Il respiro di Kussmaul è una forma di iperventilazione compensatoria. È caratterizzato da atti respiratori molto lenti e in particolare da una inspirazione

profonda e rumorosa a cui segue una breve apnea, poi una espirazione breve e gemente, infine una pausa post-espiratoria decisamente prolungata.

# Capitolo 2

## Acustica e audio digitale

### 2.1 L'acustica e il suono

L'acustica è quella branca della fisica che studia il suono, le sue cause, la sua propagazione e la sua ricezione. La percezione sonora è normalmente legata alle vibrazioni del timpano nell'orecchio. Queste vibrazioni sono provocate da piccole variazioni di pressione nell'aria. La variazione di pressione dell'aria è quindi l'equivalente fisico del suono [11]. La vibrazione provoca una successione di compressioni e rarefazioni nel mezzo dell'ambiente circostante e tale disturbo comincia a propagarsi lontano dalla sorgente in tutte le direzioni (un caso visibile è quello delle onde sull'acqua). L'effetto uditivo, consiste nella percezione da parte di un apposito dispositivo (orecchio di esseri viventi o microfoni artificiali) delle piccole e rapidissime vibrazioni emesse appunto da una sorgente sonora.

### 2.2 Onde sonore

La natura fisica del suono è di tipo ondulatorio, ovvero descrive un movimento che può essere rappresentato tramite un'onda. Si tratta di onde meccaniche che trasportano energia lontano dalla sorgente sonora. Alcune principali grandezze delle onde sono [62]:

**Periodo** Il periodo, o durata dell'oscillazione rappresenta il tempo in cui l'onda compie un'oscillazione e torna alla condizione iniziale.

**La frequenza** La frequenza rappresenta il numero di oscillazioni al secondo che l'onda compie durante la sua propagazione :

$$f = \frac{1}{t}$$

dove  $f$  è la frequenza e  $t$  la durata della singola oscillazione. L'unità di misura della frequenza è l'hertz (Hz).

**L'ampiezza** L'ampiezza invece è il valore massimo raggiunto dall'oscillazione stessa durante un periodo ed è determinata dalla quantità di energia impiegata. Viene espressa tramite i decibel.

**La forma** La forma dell'oscillazione è determinata dal numero delle componenti parziali e dal loro rapporto di frequenza, ampiezza e fase. Essa rappresenta l'aspetto dell'onda in base all'ampiezza e al tempo servendosi di coordinate cartesiane.

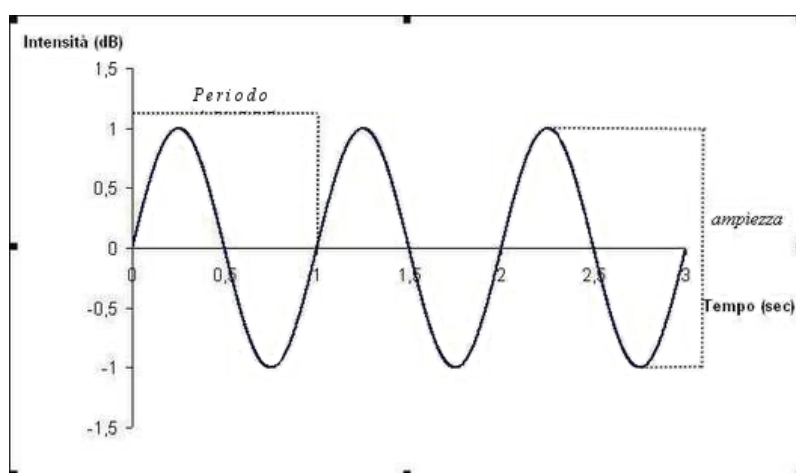


Figura 2.1: Rappresentazione delle principali grandezze dell'onda sonora [43]

## 2.3 Suono analogico e digitale

Il suono è un flusso informativo di natura temporale che scorre all'interno di apparecchiature elettroniche. Questa informazione può essere rappresentata in due forme: analogico e digitale. Una rappresentazione analogica è una rappresentazione o trasformazione di una grandezza fisica tramite una sua analogia: in altre parole, si intende un sistema in cui una quantità fisica continuamente variabile viene rappresentata da un'altra (ad esempio, la tensione di un segnale elettrico) nel modo più fedele possibile. La curva continua nel tempo delle variazioni di ampiezza viene rappresentata da una curva continua nel tempo delle variazioni di tensione elettrica.

Una rappresentazione digitale invece non cerca di imitare la curva continua di ampiezza con una curva analoga a essa, ma assegna dei numeri che rappresentano di volta in volta il valore dell'ampiezza in istanti successivi di tempo. Sarà la successione di numeri a rappresentare l'andamento della curva. La rappresentazione digitale non è continua, ma



discreta; cioè esistono degli eventi ben definiti che sono i valori dell'ampiezza in precisi istanti di tempo. I vantaggi della rappresentazione digitale e cioè di un codice simbolico sono molti: dalle operazioni di copia del segnale, a quelle di manipolazione. Ma un aspetto totalmente nuovo è la possibilità di correzione degli errori introdotti dai supporti per la memorizzazione, dove per errore si intende che alcuni numeri che rappresentano il segnale vengono letti in maniera differente da come erano stati memorizzati o trasmessi.

### 2.3.1 Discretizzazione

Per sottoporre un segnale analogico a elaborazioni digitali è necessario prima convertirlo in una sequenza di dati binari, ottenuti tramite due operazioni di discretizzazione: una discretizzazione nel dominio del tempo o campionamento che riduce gli infiniti valori di un segnale analogico in sequenze di campioni discreti e una discretizzazione nel dominio delle ampiezze o quantizzazione (del livello) che permette di rappresentare ciascun campione con un numero finito di bit. La necessità di elaborare il segnale in forma digitale porta a limitare il numero delle informazioni e quindi a misurare le grandezze solo a valori discreti del tempo, attribuendo loro un certo numero discreto di valori, variabili fra un massimo e un minimo. I segnali acquisiti diventano quindi serie di valori corrispondenti agli istanti per cui si è effettuato il campionamento.

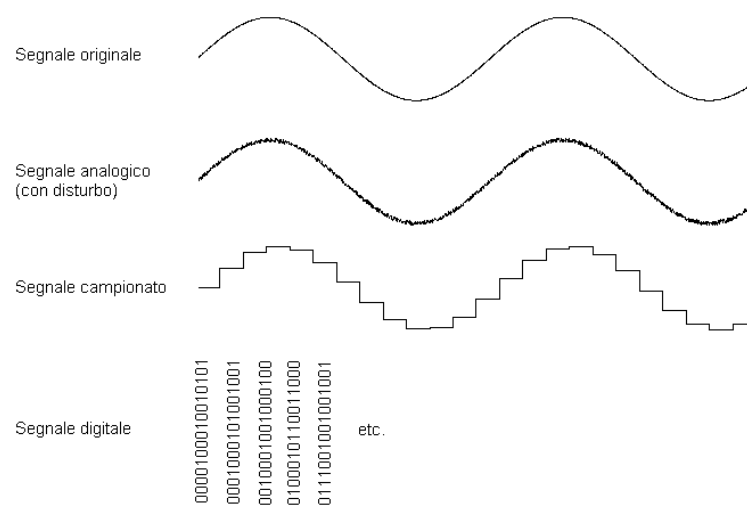


Figura 2.2: Discretizzazione di un suono da analogico a digitale [63]

### Campionamento

Campionare l'audio significa creare una sequenza di campioni, ovvero i valori di un segnale audio in diversi momenti temporali quindi ogni campione memorizzato rappresen-

ta un'ampiezza a un dato istante. Viene chiamata frequenza di campionamento il numero di volte che viene campionato il segnale in un secondo. [50].

Secondo il teorema di campionamento di Nyquist-Shannon in una conversione analogico-digitale la minima frequenza di campionamento necessaria per evitare ambiguità e perdita di informazione nella ricostruzione del segnale analogico originario, con larghezza di banda finita e nota, è pari al doppio della frequenza massima del suono che stiano convertendo. Se non viene rispettato questo teorema, cioè si ha un sottocampionamento del segnale analogico nel dominio del tempo allora nel dominio delle frequenze si ha la produzione di frequenze non proprie del segnale originario (*alias*) producendo cioè una distorsione del segnale originario divenuto ora non più fedele.

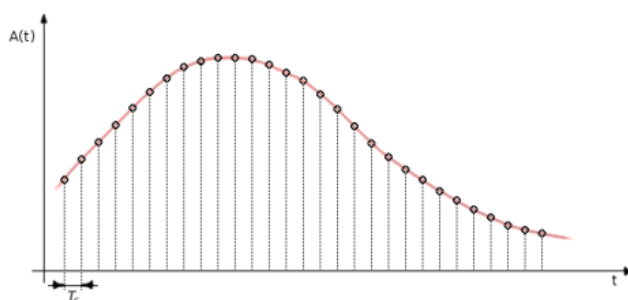


Figura 2.3: Il campionamento registra il valore dell'ampiezza a intervalli regolari prefissati [54]

## Quantizzazione

Dopo il campionamento, la conversione analogico-digitale viene completata con la quantizzazione, che consiste nell'associare a ogni campione un valore discreto.

Per ottenere ciò i valori possibili della grandezza in questione vengono innanzitutto limitati tra un massimo e un minimo intorno a dei valori discreti preventivamente definiti, definendo così le relative regioni di decisione e la dinamica del quantizzatore stesso: in tal modo il valore analogico della grandezza originaria, in corrispondenza del valore campionato in ascissa, verrà ricondotto al più prossimo dei valori discreti preventivamente definiti tramite il processo di decisione.

Con la quantizzazione vengono però introdotti degli errori detti *errori di quantizzazione* pari alla differenza tra il valore quantizzato e il suo valore reale nel campo continuo. L'ultimo passaggio per ottenere un file audio digitale consiste nella codifica, associando ai valori in uscita un numero binario. Ottenendo così una sequenza binaria che rappresenta l'onda appena discretizzata.

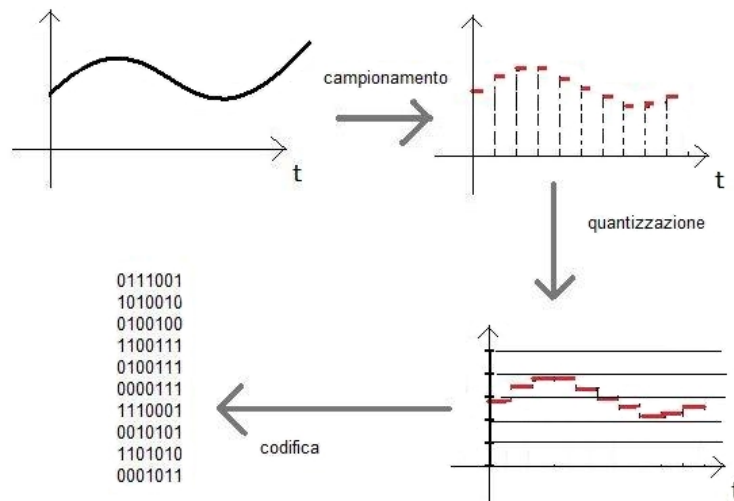


Figura 2.4: Fasi di discretizzazione di un suono [42]

## 2.4 segnale

In generale un segnale è una funzione di una o più variabili che contiene informazioni relative a un fenomeno fisico. In questo caso ci interessano i segnali sonori che sono delle funzioni dell'ampiezza rispetto al tempo. I segnali possono essere classificati secondo le seguenti proprietà:

**Continuità nel tempo** Un segnale può essere a tempo continuo o a tempo discreto a seconda che il dominio della funzione sia non numerabile o numerabile.

**Continuità nell'ampiezza** Un segnale può essere ad ampiezza continua oppure ad ampiezza discreta (o quantizzato) a seconda che l'immagine della funzione sia non numerabile o numerabile.

**Periodicità** Un segnale può essere periodico se esiste una quantità  $T$  nel dominio del tempo tale che per ogni tempo  $t$  vale  $s(T+t) = s(t)$ . Se non vale questa proprietà allora il segnale è aperiodico. Un segnale si dice quasi periodico se è composto dalla somma di segnali periodici con diverse frequenze che tra di loro stanno in rapporti non razionali.

**Determinatezza** Un segnale si dice determinato se è perfettamente noto e rappresentabile con una funzione che ne specifica l'andamento in ogni istante, invece viene chiamato aleatorio se non è completamente noto a priori, ma può assumere un qualunque andamento entro una classe di funzioni specificata da alcune caratteristiche medie dette proprietà statistiche.

**Energia** Dato un segnale  $s(t)$  definiamo l'energia del segnale come:

$$E = \int_{-\infty}^{\infty} |s(t)|^2 dt$$

La definizione acquista significato fisico quando il segnale è reale, in tal caso supponendo che  $s(t)$  rappresenti la tensione applicata o la corrente immessa a una resistenza di  $1\Omega$ , questa è l'energia da essa dissipata.

Definiamo la potenza media di un segnale continuo  $s(t)$  il limite:

$$P = \frac{1}{2T} \lim_{T \rightarrow \infty} \int_{-T}^T |s(t)|^2 dt$$

Il segnale  $s(t)$  è detto a energia finita se  $E$  è finito e diverso da zero. Il segnale è detto a potenza media finita se  $P$  è finito e diverso da zero. Una classe importante di segnali a potenza media finita è costituita dai segnali periodici; in tal caso l'energia è finita e la potenza media coincide con quella calcolata in un periodo.

**Envelope di un segnale** Dato un segnale nel dominio del tempo, diciamo che la sua envelope è una curva che varia molto meno rapidamente del segnale e che traccia il contorno di esso [15]. La figura 2.4 illustra due envelope, una superiore e una inferiore di un segnale sinusoidale.

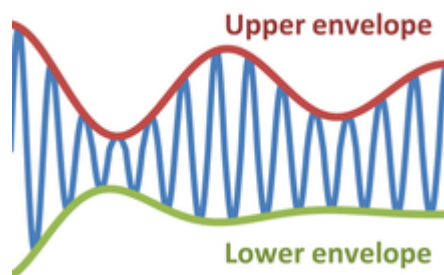


Figura 2.5: Envelope superiore e inferiore di una onda sinusoidale [58]

## 2.5 Analisi nel dominio del tempo e nel dominio della frequenza

I principali metodi di analisi del segnale possono essere riassunti nei concetti di analisi nel dominio del tempo e analisi nel dominio della frequenza. È importante osservare che questi due modi di affrontare un problema sono tra loro intercambiabili, nel senso

che, sotto opportune condizioni, nessuna informazione viene persa nel passare da un dominio all'altro. Il vantaggio che deriva dall'introduzione dei due domini è la possibilità di cambiare la prospettiva con la quale si osserva un dato fenomeno. In questo modo un problema che appare di difficile soluzione in un dominio può risultare molto più semplice nell'altro [41].

**Analisi nel dominio del tempo** Questa forma di rappresentazione è quella che ci è maggiormente familiare; in essa appaiono le variazioni subite dal segnale al trascorrere del tempo. Lo strumento più frequentemente usato e che opera notoriamente nel dominio del tempo è l'oscilloscopio.

**Analisi nel dominio della frequenza** Invece di analizzare le variazioni del segnale al passare del tempo, mostra come e quanto un segnale si suddivide o è distribuito nelle varie bande di frequenza, definite all'interno di un dato range. Viene chiamata analisi spettrale e lo strumento matematico che consente di trasferire lo studio dei segnali e dei sistemi dal dominio del tempo al dominio della frequenza è la trasformata di Fourier.

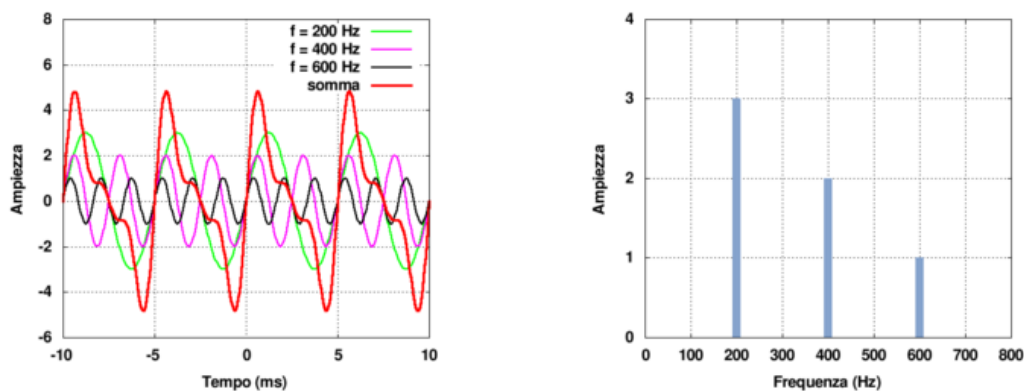


Figura 2.6: Esempi di grafici nel dominio del tempo e nel dominio delle frequenze [46]

## 2.6 Analisi armonica e trasformata di Fourier

L'analisi armonica è la branca della matematica che studia la rappresentazione delle funzioni o dei segnali come sovrapposizione di onde fondamentali. Indaga e generalizza la nozione di serie di Fourier e trasformata di Fourier.

L'interpretazione dello sviluppo in serie di Fourier è che un segnale periodico di potenza finita si può sviluppare come combinazione lineare di funzioni periodiche semplici la cui frequenza è un numero intero. Il tutto si generalizza facilmente al caso in cui il periodo sia  $T$  e quindi la frequenza fondamentale abbia il valore  $1/T$ .

L'importanza della formula è che tutta l'informazione di una funzione periodica continua può essere espressa con un'infinità numerabile (quindi discreta) di valori complessi.

### 2.6.1 Serie di Fourier

La serie di Fourier è la rappresentazione di una funzione mediante una combinazione lineare di funzioni sinusoidali fondamentali. Un segnale  $f(t)$  è una funzione e quindi lo si può scomporre usando una serie di Fourier. La scomposizione di un segnale attraverso una serie di Fourier ci dà una descrizione delle frequenze che lo compongono [6] [47].

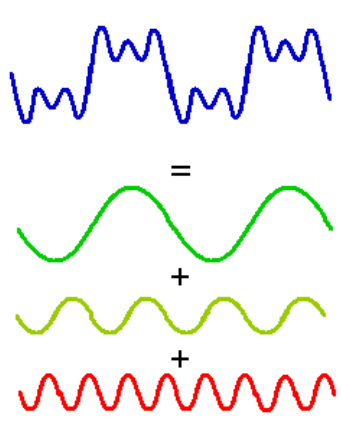


Figura 2.7: Scomposizione di un segnale [44]

Sia  $f(x)$  una funzione definita sui numeri reali e periodica di periodo  $2\pi$ . Diremo che  $f(x)$  è sviluppabile in serie di Fourier se esistono dei coefficienti  $a_0, a_k, b_k$  con  $k$  numero naturale tali che

$$f(x) = a_0 + \sum_{k \in \mathbb{N}} (a_k \cos(kx) + b_k \sin(kx))$$

La formula precedente è lo sviluppo in serie di Fourier della funzione  $f$  e i coefficienti di Fourier si possono calcolare nel seguente modo [23]:

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx \quad a_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx \quad b_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx$$

### 2.6.2 Trasformata di Fourier

La trasformata di Fourier permette di scomporre e successivamente ricombinare, un segnale generico in una somma infinita di sinusoidi con frequenze, ampiezze e fasi diverse. L'insieme di valori in funzione della frequenza è detto spettro di ampiezza e spettro di fase. Se il segnale in oggetto è un segnale periodico, la sua trasformata di Fourier è un insieme di valori discreti, che in tal caso prende il nome di spettro discreto. Mentre nel caso in cui il segnale sia non periodico lo spettro è continuo, e tanto più è esteso lungo l'asse

delle frequenze quanto più è limitato nel dominio originario della variabile indipendente, e viceversa. Se il segnale ha un valore medio diverso da zero la serie restituisce anche una componente costante che lo rappresenta.

Sia  $f \in L^1(\mathbb{R}^n)$  una funzione integrabile, la trasformata continua di Fourier, detta anche semplicemente trasformata di Fourier, è definita nel seguente modo [6][47]:

$$\mathbb{F}(f) = t \mapsto \frac{1}{(2\pi)^{\frac{n}{2}}} \int_{\mathbb{R}^n} e^{-i \cdot t \cdot x} dx$$

### 2.6.3 Trasformata discreta di Fourier

La trasformata di Fourier opera su funzioni continue, sia nel dominio del tempo che delle frequenze. Al contrario, la trasformata discreta di Fourier opera su funzioni a dominio discreto. Sia  $S_n$  un insieme di sequenze periodiche di periodo  $n$ . Un generico elemento di  $S_n$  si può scrivere come  $y = \{y_j\}_{j \in \mathbb{N}}$  e si può pensare come un segnale periodico discreto nel quale  $y_j$  è il valore del segnale al tempo  $j$ . Definiamo la trasformata discreta di Fourier di  $y$  come la sequenza

$$(\mathbb{F}_n(y))_k = z_k$$

dove

$$z_k = \sum_{j=0}^{n-1} y_j \bar{w}^{j \cdot k} \quad w = \exp\left(\frac{s\pi i}{n}\right)$$

[6]

## 2.7 Wavelet

Tramite la trasformata di Fourier il passaggio dal dominio del tempo al dominio delle frequenze comporta una perdita totale dell'informazione temporale. Viene introdotto il concetto di scala sostituendolo a quello delle frequenze: invece di spostarsi da una frequenza all'altra ci si sposta da una risoluzione più fine a una più grossolana. Una wavelet è una forma d'onda oscillante di lunghezza finita, dove assume valori sia positivi che negativi. La differenza principale tra la trasformata di Fourier e le wavelet è che le wavelet sono localizzate sia nel tempo che nella frequenza mentre la trasformata di Fourier standard è localizzata solo in frequenza.

Un esempio di wavelet è rappresentato dalla figura seguente:

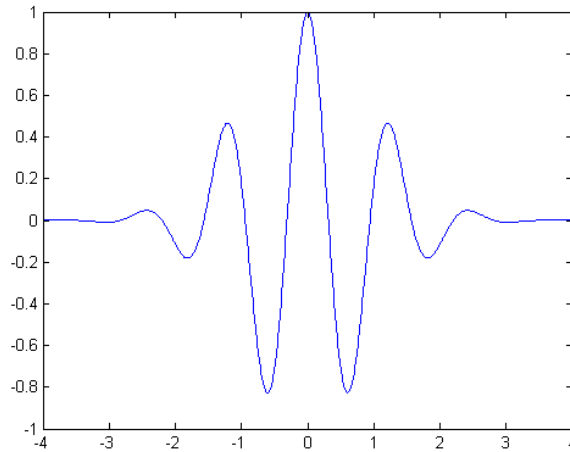


Figura 2.8: Esempio di una wavelet [40]

Una funzione  $\psi(t)$  è una wavelet se soddisfa i seguenti criteri:

- Una wavelet deve avere energia finita.

$$E = \int_{-\infty}^{\infty} |\psi(t)|^2 dt < \infty$$

- Se  $\Psi$  è la trasformata di Fourier della wavelet  $\psi(t)$  allora  $\Psi(0) = 0$ . Cioè la wavelet non ha componenti nella frequenza zero.
- La trasformata di Fourier di una wavelet complessa deve restituire un valore reale e deve essere nulla nelle frequenze negative.

### 2.7.1 Trasformata wavelet

Applicando la trasformata wavelet a un segnale si ottiene una matrice contenenti valori che indicano l'intensità di una certa frequenza in un istante di tempo discreto. Per ottenere questa matrice i principali algoritmi eseguono il prodotto tra la funzione di ingresso e la wavelet al tempo 0. La wavelet viene poi traslata a un tempo  $b$  e viene eseguito nuovamente il prodotto tra essa e la funzione in ingresso, questa operazione si ripete fino ad arrivare alla fine della funzione in ingresso [28].

A questo punto si è riempita la prima riga della matrice contenente le informazioni. Per inserire una seconda riga alla matrice vengono ripetuti i passaggi precedenti applicando la wavelet con una scala maggiore, cioè usando la wavelet allungata nell'asse temporale. Riempendo così la matrice si ha che ogni riga indica la scala mentre ogni colonna rappresenta il tempo.



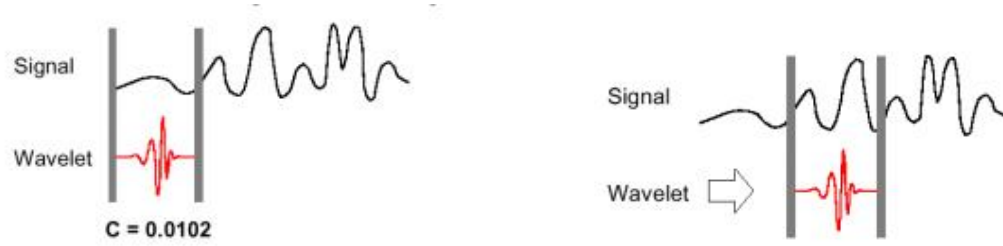


Figura 2.9: Applicazioni della wavelet nel segnale [52]

### Trasformata wavelet continua e discreta

Sia  $\psi(t)$  una wavelet, definiamo la trasformata wavelet continua di una funzione  $x(t)$  come

$$X(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) x(t) dt$$

La scala o parametro di dilatazione  $a$  corrisponde all'informazione di frequenza e il parametro di traslazione  $b$  si riferisce alla locazione della funzione wavelet mentre viene fatta spostare lungo il segnale, quindi corrisponde all'informazione temporale nella trasformata. Esiste anche la trasformata wavelet discreta che opera su segnali a tempo discreto e che restituisce coefficienti discreti nel dominio delle wavelet. I parametri  $a$  e  $b$  assumono valori in una griglia discreta [10][24] :

$$a = 2^{-j} \quad b = k \cdot 2^{-j} \quad j, k \in \mathbb{Z}$$



# Capitolo 3

## Filtri digitali

### 3.1 Denoising

Uno degli ostacoli principali dell'analisi computerizzata dei suoni polmonari è la presenza di rumore nei segnali. In questo caso per rumore si intendono sia i suoni provenienti da fattori estremi come ventilatore, aria condizionata, e altri rumori ambientali che possono contaminare i segnali sonori del polmone, sia dai suoni interni emessi dal paziente come il battito cardiaco. La natura rumorosa dei suoni polmonari è un fattore di impedimento che vieta l'identificazione di funzioni utili per la diagnostica. Quindi il denoising di segnali sonori polmonari è d'obbligo per una efficace diagnosi e in questo capitolo approfondiremo le varie tecniche per l'eliminazione dei rumori.

### 3.2 Classificazione dei filtri digitali

Un filtro digitale è una funzione in grado di modificare il contenuto armonico di un segnale elettrico complesso. Queste modifiche si esprimono in termini di fasci di frequenze che vengono evidenziate o sopresse a seconda delle caratteristiche del tipo di filtro usato. Una importante penalizzazione per i filtri digitali è che sono limitati nelle frequenze: un filtro digitale deve sempre e comunque rispettare il teorema di Nyquist altrimenti si ottengono segnali disturbati da aliasing.

Si distinguono due classi di filtri: lineari e non lineari.

**Filtri lineari** La funzione dà in uscita un valore che è una combinazione lineare dei valori compresi nella finestra del segnale appena analizzato, cioè l'output del filtro può mettersi in relazione con i valori presi in ingresso. Si possono definire filtri lineari quelli che puliscono il segnale dal rumore oppure che esaltano le discontinuità.

**Filtri non lineari** Per questo tipo di filtro non è possibile definire un operatore lineare; solitamente sono operatori di rango, cioè operatori che agiscono sui valori presi in ingresso dopo averli ordinati quindi l'output dipende contemporaneamente da più ingressi.

La differenza sostanziale tra i due tipi di filtri è che, mentre per i primi si può applicare la trasformata di Fourier con tutte le sue proprietà, nei secondi questa operazione non è possibile.

### 3.2.1 Categorie di filtri

I filtri si dividono sostanzialmente in quattro principali categorie [62]:

#### Filtro passa alto (high pass)

Il filtro passa alto permette il passaggio di una banda di frequenze al di sopra di una determinata frequenza limite (*frequenza di taglio*). Il taglio del filtro indica la zona della banda di frequenze che divide le due sezioni: frequenze passanti (alte), frequenze tagliate (basse).

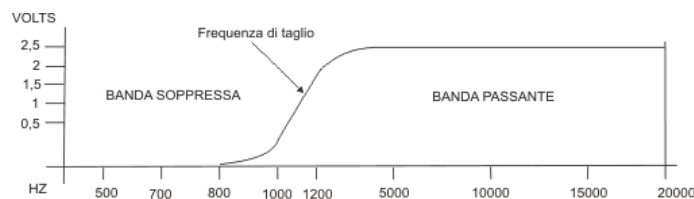


Figura 3.1: Esempio di un filtro passa alto[62]

#### Filtro passa basso (low pass)

Il filtro passa basso è esattamente l'opposto del filtro passa alto cioè permette il passaggio di una banda di frequenze al di sotto di una determinata frequenza limite.

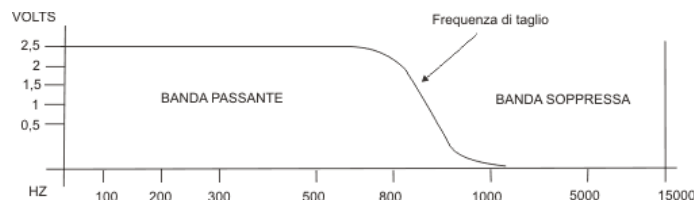


Figura 3.2: Esempio di un filtro passa basso [62]

### Filtro passa banda (band pass)

Il passa banda deriva dall'accoppiamento di un passa alto con un passa basso, quindi lascia passare solo una banda ristretta di frequenze che rientrano tra le frequenze di taglio dei due filtri accoppiati. Comunque bisogna tener conto che al di sopra e al di sotto di tali frequenze le altre componenti del segnale vengono eliminate, non in maniera netta, ma in modo graduale dipendente dalla pendenza della curva di selettività del filtro che varia a seconda del livello qualitativo del filtro stesso.

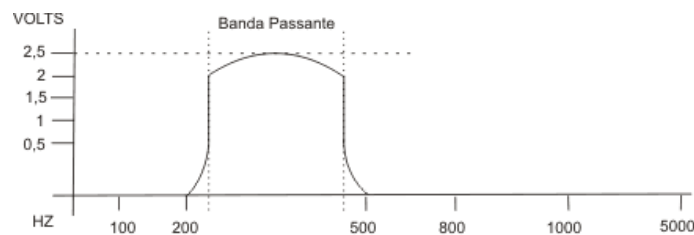


Figura 3.3: Esempio di un filtro passa banda [62]

### Filtro soppressore di banda (band reject)

Il filtro a soppressione di banda è esattamente il contrario del filtro passa banda, in quanto pur essendo costituito dall'unione di un passa alto con un passa basso, respinge solo una ristretta banda di frequenze compresa tra le due frequenze di taglio lasciando passare tutte le altre.

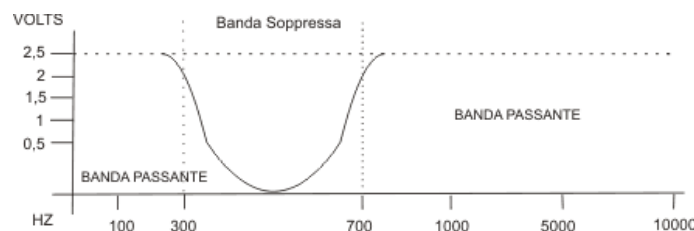


Figura 3.4: Esempio di un filtro soppressore di banda [62]

## 3.3 Filtrare i suoni respiratori

Entrando più nel dettaglio, durante l'analisi della letteratura attuale abbiamo riscontrato diversi tipi di filtri per l'eliminazione del rumore (denoising) e per eliminare il battito cardiaco. L'eliminazione dei suoni emessi dal cuore è un problema cruciale per il riconoscimento del respiro. Nello scenario previsto abbiamo un file audio del respiro preso

grazie a uno stetoscopio elettronico; molti modelli di stetoscopi già consentono funzioni per filtrare il battito cardiaco [59], rilasciando quindi un suono a cui bisogna applicare solo un filtraggio per il denoising, però nell'eventualità in cui uno stetoscopio non sia fornito di queste funzioni bisogna trovare una tecnica di filtraggio ottimale che la maggior parte delle volte consiste nell'applicare più di un filtro al suono che stiamo analizzando. Si può subito notare attraverso l'analisi della letteratura approfondita nel capitolo 5, che il filtro usato maggiormente è quello passa banda, che permette di eliminare le frequenze che non rientrano nel range di frequenza dei suoni respiratori.

Per una migliore comprensione dell'analisi dei sistemi dei capitoli successivi si andranno a illustrare alcune tecniche da loro usate.

### **Downsampling**

Un filtro di downsampling o di sottocampionamento riduce la frequenza di campionamento di un segnale. Di solito lo scopo di questo filtro è di ridurre la complessità computazionale dei successivi trattamenti del segnale. Il fattore di sottocampionamento è un numero intero o razionale maggiore di uno. Questo fattore moltiplica il tempo di campionamento o equivalentemente divide la frequenza di campionamento. Una implementazione di un filtro di sottocampionamento deve fare attenzione a non violare le condizioni del teorema sul campionamento di Shannon-Nyquist, altrimenti ci sarà la presenza di aliasing nel segnale di output. Questa tecnica viene usata sia da [36] che da [2].

### **Filtro mediano**

Il filtro mediano è un filtro non lineare ed è usato per la rimozione dei picchi di rumore dal segnale e influenzano di solito solo una piccola percentuale dei campioni ma in modo notevole. La tecnica di questo filtro è quella di rimpiazzare i campioni del segnale con la mediana dei campioni calcolata in un certo intervallo di campioni. Questo intervallo si chiama finestra, la quale si muove su tutto il segnale campione per campione [45]. Questo filtro viene usato da [2].

### **Funzione finestra**

Una finestra  $w(n)$  è un segnale che, nel dominio del tempo, viene moltiplicato per il segnale da analizzare  $x(n)$  e rappresenta una funzione che preso un intervallo pone a zero i valori esterni a esso. Stabilendo la dimensione della finestra, il segnale viene segmentato e per ogni segmento viene applicata la funzione. Quando un segnale o un suo segmento

viene moltiplicato per una funzione finestra, anche il suo risultato assume valori nulli fuori dall'intervallo, o più in generale che si annulli in maniera sufficientemente rapida.

Da notare che, nonostante l'applicazione della finestra, non si procede al taglio del solo segnale compreso all'interno di essa, ma si mantiene l'intera onda sonora che però è in gran parte ridotta al minimo suono in quanto esterna alla finestra opportunamente posizionata sull'onda. In tal modo si mantiene la corretta informazione sull'andamento reale del suono. Si procede così per ogni segmento all'applicazione della funzione finestra.

La finestatura di un segnale ha l'importante conseguenza di causare la dispersione spettrale e la scelta della finestra viene fatta in modo tale che la trasformata della finestra si riduca il più possibile a un impulso, cioè il lobo principale sia il più stretto possibile e i lobi secondari siano i più bassi possibili.

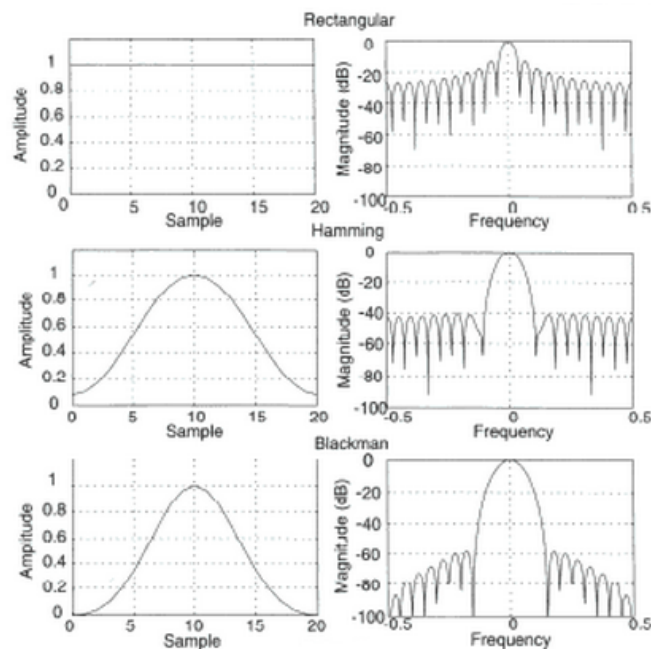


Figura 3.5: Funzioni finestra [62]

Considerando  $N$  come l'ampiezza in numero di campioni di una finestra a tempo discreto tipicamente rappresentato con una potenza di 2 e  $n$  un numero intero, che assume valori da 0 a  $N - 1$ .

Ci sono varie funzioni finestre, per esempio:

- La finestra rettangolare, tronca il segnale ponendo un valore unitario in tutto il tratto in cui non è nulla.

$$w(n) = 1$$

- La finestra di Hamming descritta dall'equazione

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right)$$

è stata progettata per minimizzare il livello del lobo laterale, mentre mantiene approssimativamente la stessa larghezza del lobo principale, e viene usata da [2].

- La finestra di Hann o Hanning descritta dall'equazione:

$$w(n) = 0.5 \cos\left(1 - \frac{2\pi n}{N-1}\right)$$

Questa funzione viene usata da [26] e da [20].

- La finestra di Blackman descritta dall'equazione:

$$w(n) = \frac{1-0.16}{2} - \frac{1}{2} \cos\left(1 - \frac{2\pi n}{N-1}\right) + \frac{0.16}{2} \cos\left(\frac{4\pi n}{N-1}\right)$$

Restituisce il lobo principale molto più ampio rispetto alle funzioni finestra appena viste e viene usato da [2].



# Capitolo 4

## Metodi di base

In questo capitolo si introducono le premesse teoriche necessarie per la comprensione di:

- Algoritmi usati nello stato dell'arte.
- Algoritmi utili per il sistema suggerito in conclusione.
- Algoritmi dei quali è stata valutata l'adeguatezza alla soluzione del problema.

Oltre alla descrizione degli algoritmi che si vedranno nell'analisi della letteratura si è data una particolare importanza nella descrizione di algoritmi utili per la soluzione del problema. Dopo un'analisi delle possibili tecniche risolutive si è scelto di suggerire un sistema che utilizza un algoritmo di beat detection insieme a un algoritmo di clustering. In più vi è una descrizione sulle reti neurali perché ritenuta un'ulteriore possibile soluzione.

### 4.1 Beat detection e algoritmi di onset detection

Con il termine beat detection intendiamo il riconoscimento del ritmo o del tempo di una musica a partire da una sua codifica digitale. Ad un livello intuitivo siamo tutti in grado di capire cos'è il ritmo e siamo anche in grado di capire qual'è il ritmo di una canzone quando la si ascolta.

Il ritmo è definito da una successione di accenti, intendendo con accento il maggior rilievo (variazione di intensità, altezza o timbro) che alcuni suoni hanno rispetto ad altri nell'ambito di un brano o una frase musicale. La sequenza degli accenti di un brano musicale tende normalmente a ripetersi a intervalli regolari ed è questa ripetizione che viene chiamata ritmo del brano: la più breve sequenza non periodica (quella che viene ripetuta) viene anche chiamata cellula ritmica.

Possiamo definire un onset come un accento importante alla caratterizzazione del ritmo, cioè l'istante in cui la potenza del segnale raggiunge un picco di maggiore energia. Per trovare il ritmo quindi dobbiamo prima passare dagli onsets. Per la determinazione degli onsets il segnale viene segmentato in frammenti delimitati da due eventi successivi che rappresentano un qualsiasi fenomeno che alteri l'andamento del suono che si sta esaminando. La scelta della dimensione dei segmenti può variare molto e dipende dal tipo di ricerca che si vuole effettuare.

La maggior parte degli algoritmi di onset detection seguono uno schema assimilabile a quello illustrato nella figura 4.1 e descritto da *Bello et al.* [4].

### 4.1.1 Preprocessing

Il concetto di preprocessing implica la trasformazione del segnale originale con lo scopo di accentuare o attenuare alcuni aspetti del segnale in funzione delle operazioni successive. Questa fase è opzionale e ci sono vari modi per eseguire queste trasformazioni come per esempio la separazione nelle frequenze e in questo caso l'informazione viene separata in bande di frequenza.

### 4.1.2 Reduction

La riduzione è il processo di trasformare il segnale audio in un segnale con una frequenza di campionamento ridotta e che manifesta in modo più evidente la posizione degli onsets. Questo è il cuore degli algoritmi di onset detection. Chiamiamo l'output di questa fase anche *detection function*. Classifichiamo i metodi di riduzione in due categorie: metodi basati su caratteristiche esplicite del segnale e metodi basati su modelli probabilistici.

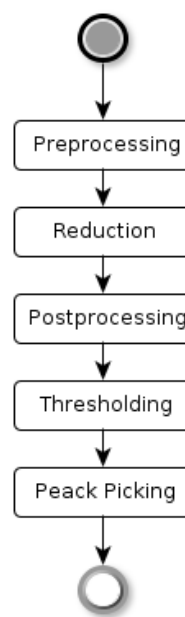


Figura 4.1: Diagramma di attività di un generico algoritmo di beat detection

**Metodi basati su caratteristiche esplicite del segnale.** Le caratteristiche esplicite di un segnale possono rientrare nelle categorie: temporali, spettrali, spettrali di fase e spet-trotemporali. Per esempio analizzando il dominio del tempo del segnale si nota spesso che una occorrenza di un onset di solito è accompagnata da un aumento dell'ampiezza del segnale. Alcuni metodi di onset detection si avvantaggiano di questa proprietà e creano

una detection function che segue l'envelope del segnale. Questi metodi danno risultati soddisfacenti nei casi in cui c'è un onset molto forte rispetto al sottofondo.

**Metodi basati su modelli probabilistici.** I metodi statici per l'onset detection sono basati sull'assunto che il segnale può essere descritto da qualche modello probabilistico. Quindi si può costruire un sistema che, attraverso l'inferenza statistica, stabilisce gli istanti in cui si hanno con maggiore probabilità variazioni brusche del segnale in base alle osservazioni disponibili. Il successo di questa tecnica dipende dall'aderenza del modello alla reale distribuzione dei dati e può essere quantificato usando misure di verosimiglianza o criteri di selezione del modello Bayesiano, dove il modello Bayesiano si rappresenta tramite un grafico probabilistico contenente un insieme di variabili stocastiche e le dipendenze tra loro.

### 4.1.3 Postprocessing

Questa fase è opzionale e serve per facilitare le fasi successive attraverso l'aumento dell'uniformità e della consistenza di alcuni aspetti della detection function che vengono trasformati in massimi locali isolati e facili da trovare. In questa fase possiamo trovare ad esempio: algoritmi di riduzione del rumore e algoritmi di selezione di parametri utili al calcolo del thresholding nelle fasi successive.

### 4.1.4 Thresholding

Ci possono essere dei picchi nella detection function che non sono correlati ad un onset di interesse. Quindi è necessario definire un valore di soglia che separi in modo efficace i picchi che corrispondono ad un evento di interesse e quelli che non vi corrispondono. In questa fase un onset che corrisponde ad un evento di interesse è definito come un picco nel quale la detection function supera una certa soglia. Ci sono due approcci principali alla scelta della soglia:

**Thresholding fisso** La soglia in questione non varia nel corso della vita dell'algoritmo.

**Thresholding adattativo** La soglia in questione varia ed è di solito una certa funzione della detection function. Ad esempio:

**Linear smoothing** La soglia è una combinazione lineare degli elementi all'interno della finestra corrente della detection function.

**Non linear smoothing** La soglia è una combinazione non lineare degli elementi all'interno della finestra corrente della detection function. Ad esempio una combinazione dei quadrati.

**Percentiles smoothing** I metodi precedenti hanno lo svantaggio che quando ci sono dei picchi relativamente grandi seguiti da picchi piccoli, questi ultimi vengono nascosti. Per evitare questo problema si possono usare metodi basati sui percentili come per esempio la mediana.

### 4.1.5 Peak picking

Se la function detection è stata ben progettata gli onset si rifletteranno in una ben precisa caratteristica della funzione di rilevamento e saranno facilmente individuati. Nella fase di scelta dei picchi quindi si devono solamente identificare i massimi locali al di sopra del threshold per ottenere gli onset del segnale.

## 4.2 Clustering

Il clustering è un insieme di tecniche di analisi atte a trovare informazioni strutturali all'interno di un insieme di dati. I meccanismi di clustering separano e organizzano dati non etichettati in vari gruppi o cluster, ognuno dei quali contiene dati che sono simili tra loro. Un buon algoritmo di clustering produce cluster nel quale la somiglianza tra elementi dello stesso cluster (somiglianza *intracluster*) è alta, mentre risulta bassa quella tra elementi appartenenti a cluster diversi (somiglianza *intercluster*). Per una discussione approfondita sul clustering si rimanda a [34].

Se vogliamo essere più formali possiamo dire che un  $m$ -clustering di un insieme di dati  $X$  rispetto a una distanza o misura di dissimilarità  $d : P(X) \rightarrow \mathbb{R}$  e a una soglia  $t$  è  $C_1, \dots, C_n$  tale che:

**partizione**  $C_1, \dots, C_n$  è una partizione di  $X$  cioè: ogni  $C_i$  è un sottoinsieme di  $X$ ; i  $C_i$  sono a due a due disgiunti; l'unione dei  $C_i$  da  $X$

**intersimilarità** per ogni  $i$ :  $d(C_i) < t$

### 4.2.1 Algoritmi di clustering

Di seguito descriveremo alcuni tipi di algoritmi di clustering, ci sono principalmente due categorie, una prima categoria dipende dalla possibilità che un elemento possa essere assegnato a più cluster [39]:

**Clustering esclusivo** dove ogni elemento può appartenere a un solo cluster.

**Cluster non-esclusivo** dove un elemento può appartenere a più cluster con gradi di appartenenza diversi.

Una seconda suddivisione delle tecniche di clustering tiene conto del tipo di algoritmo usato per dividere lo spazio:

**Clustering partizionale** in cui per definire l'appartenenza a un gruppo viene utilizzata una distanza da un punto rappresentativo del cluster, avendo prefissato il numero di gruppi della partizione risultato. Un esempio di questa suddivisione è il clustering sequenziale: Questi algoritmi leggono gli elementi di input uno per volta in sequenza. In questo tipo di algoritmi non si conoscono a priori il numero di cluster presenti nell'input. Viene presa in input una soglia  $\theta$  e il massimo numero di cluster permessi  $q$ . La tecnica di questi algoritmi consiste nel leggere un elemento e assegnarlo o a un cluster già esistente o a un nuovo cluster a seconda della distanza tra l'elemento e i cluster esistenti. Uno schema generale per implementare un algoritmo di clustering sequenziale è dato nell'algoritmo 1.

<p><b>input</b> : A sequence of data <math>x_1, \dots, x_n</math>  <b>output</b>: A clustering of the data <math>C_1, \dots, C_m</math></p> <pre> 1 <math>m \leftarrow 1</math> 2 <math>C_m \leftarrow \{x_1\}</math> 3 <b>for</b> <math>i \leftarrow 2</math> <b>to</b> <math>n</math> <b>do</b> 4   find <math>k</math> such that the distance between <math>x_i</math> and <math>C_k</math> is minimum 5   <b>if</b> <math>\text{distance}(x_i, C_k) &gt; \text{threshold}</math> <b>then</b> 6     increments <math>m</math> 7     <math>C_m \leftarrow \{x_i\}</math> 8   <b>else</b> 9     <math>C_m \leftarrow C_m \cup \{x_i\}</math> </pre>
--

**Algorithm 1:** Sequential clustering algorithm

Nel caso in cui non si conosca in anticipo neanche il valore di soglia allora si può stimare il numero di cluster che ci sono nell'input usando l'algoritmo 2.

<p><b>input</b> : A sequence of data <math>x_1, \dots, x_n</math>  <b>output</b>: An estimation of the number of clusters</p> <pre> 1 <b>for</b> <math>\text{threshold} \leftarrow a</math> <b>to</b> <math>b</math> <b>step</b> <math>c</math> <b>do</b> 2   run clustering algorithm with current threshold 3   store the number of clusters obtained 4 <b>return</b> the most frequent number of clusters </pre>
---

**Algorithm 2:** Estimation of the number of clusters

**Clustering Gerarchico** in cui viene costruita una gerarchia di partizioni caratterizzate da un numero crescente o decrescente di gruppi, visualizzabile mediante una rappre-

sentazione ad albero, in cui sono rappresentati i passi di accorpamento e divisione dei gruppi. Le strategie per il clustering gerarchico sono tipicamente di due tipi:

**Agglomerativo** Questo tipo di algoritmo produce una sequenza di clustering che hanno numero di cluster decrescente: inizialmente assume che ogni cluster contenga un singolo elemento, ad ogni passo vengono uniti i cluster tra loro più vicini fino ad ottenere pochi cluster tra loro differenti. Questo tipo di clustering necessita di tecniche per valutare la somiglianza tra cluster.

**Divisivo** questo tipo di algoritmo produce una sequenza di clustering che hanno numero di cluster crescente. A ogni iterazione un algoritmo di questo tipo prende i cluster risultati dall'iterazione precedente e li divide.

## 4.3 Reti neurali

Le *reti neurali artificiali* sono modelli matematici e computazionali della corteccia cerebrale. Gli elementi di base di una rete neurale artificiale sono:

**Neuroni artificiali o unità** Sono dei costrutti matematici che in qualche misura imitano le proprietà dei neuroni viventi.

**Collegamenti** I neuroni artificiali sono connessi attraverso *collegamenti*. Ogni collegamento ha un *peso* numerico, in particolare il collegamento tra l'unità  $i$  e l'unità  $j$  ha peso  $w_{i,j}$ .

Una certa unità  $i$  può inviare un *segnale* o *attivazione*  $a_i$  su tutti i suoi collegamenti. Il peso di un collegamento determina la forza e il segno del segnale. Assumiamo l'esistenza di una certa unità di input che manda il segnale  $a_0$ . Ogni unità si comporta nel modo seguente:

1. Calcola la somma pesata dei suoi input:

$$in_j = \sum_{i=0}^n w_{i,j} a_i$$

2. Calcola il proprio output in base al valore calcolato al punto precedente e ad una *funzione di attivazione*  $g$ :

$$a_j = g(in_j)$$

La funzione di attivazione di solito è una funzione del tipo:

$$g(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0 \end{cases} \quad \text{o} \quad g(x) = \frac{1}{1+e^{-x}}$$

Ci sono due tipi di reti neurali:

**Feed forward** i collegamenti tra unità sono unidirezionali. In questo caso la rete neurale è un grafo diretto aciclico. Una rete neurale di tipo feed forward rappresenta una funzione dei suoi input nel senso che non ha uno stato interno o memoria. Una rete feed forward è di solito divisa in *layers* o livelli. Ogni livello è un insieme di unità, ogni unità riceve input solo da unità del livello immediatamente precedente. Una rete con un solo livello si dice anche *single-layer* o *mono-livello*.

**Recurrent network** ci possono essere dei cicli nei collegamenti. Una rete di questo tipo rappresenta un sistema dinamico che può anche oscillare o avere un comportamento caotico. Inoltre la risposta di una recurrent network dipende dallo stato iniziale della rete e dall'input quindi questo tipo di rete può supportare una memoria a breve termine.

## 4.4 Dimensione frattale

In geometria frattale la dimensione frattale è una quantità statistica che da una indicazione di quanto completo appare un frattale per riempire lo spazio. La dimensione frattale inoltre è una misura della complessità di un insieme di dati. Viene usata per analizzare segnali in un vasto numero di ricerche scientifiche. Una proprietà della dimensione frattale consiste nell'essere indipendente dal contenuto di energia nel segnale ma è dipendente dalle frequenze. Ci sono molti modi di definire matematicamente la dimensione frattale ad esempio quello illustrato in [22].

## 4.5 Catene di triple Markoviane

### 4.5.1 Nozioni di base di probabilità

**Evento, Spazio Campionario** Considerando un esperimento il cui risultato non è prevedibile con certezza, si definisce un evento elementare  $\omega$  come uno dei possibili esiti dell'esperimento stesso e si definisce un evento come un qualsiasi insieme di eventi elementari.

Viene definito spazio campionario, solitamente indicato con il nome  $\Omega$ , l'insieme di tutti i possibili esiti dell'esperimento, cioè l'insieme di tutti gli eventi elementari.

**$\sigma$ –algebra** Dato un insieme  $\Omega$ , si definisce  $\sigma$ –algebra su  $\Omega$  una famiglia  $F$  di sottoinsiemi di  $\Omega$  tale che:

- $\Omega$  appartiene ad  $F$ .
- L'insieme  $F$  è chiuso rispetto al complementare.
- L'insieme  $F$  è chiuso rispetto ad unione numerabile.

**Spazio misurabile, Insiemi misurabili** Uno spazio misurabile è una coppia  $(\Omega, F)$  tale che  $F$  è una  $\sigma$ –algebra su  $\Omega$ . Gli elementi di  $F$  sono detti insiemi misurabili in  $\Omega$ .

**Topologia, Spazio topologico** Dato un insieme  $X$ , una topologia su  $X$  è un insieme  $T$  di sottoinsiemi di  $X$  tale che:

- L'insieme vuoto e  $X$  appartengono a  $T$ .
- L'unione di una quantità arbitraria di insiemi appartenenti a  $T$  appartiene a  $T$ .
- L'intersezione di un numero finito di insiemi appartenenti a  $T$  appartiene a  $T$ .

Uno spazio topologico è una coppia  $(X, T)$ , dove  $X$  è un insieme e  $T$  è una topologia su  $X$ . In uno spazio topologico gli insiemi che costituiscono  $T$  si dicono aperti in  $X$ .

**Funzione misurabile** Sia  $(X, F)$  uno spazio misurabile e  $(Y, T)$  uno spazio topologico. Una funzione  $f : X \rightarrow Y$  viene detta misurabile se la controimmagine di ogni elemento di  $T$  è in  $F$ .

**Variabile casuale o aleatoria** Sia dato uno spazio campionario  $\Omega$  su cui è definita una misura di probabilità  $\nu$ , una variabile casuale è una funzione misurabile dallo spazio campionario a uno spazio misurabile.

**Processo stocastico** Si definisce processo stocastico una famiglia di variabili aleatorie

$$\{X(t) \mid t \in T \subseteq \mathbf{R}^+\}$$

dipendenti dal tempo, definite su un unico spazio campione  $\Omega$  finito e che assumono valori in un insieme definito spazio degli stati del processo. Un processo stocastico è quindi un insieme di funzioni che evolvono nel tempo, ognuna delle quali è associata ad un



determinato elemento dello spazio campione, così che il risultato di un esperimento casuale corrisponde di fatto all'estrazione di una di queste funzioni. Un processo stocastico discreto è un processo stocastico nel quale  $T$  è un sottoinsieme dei numeri naturali.

### 4.5.2 Processo di Markov

Un processo stocastico markoviano o processo di Markov è un processo stocastico nel quale la probabilità di transizione che determina il passaggio a uno stato di sistema dipende unicamente dallo stato di sistema immediatamente precedente (proprietà di Markov) e non dal come si è giunti a tale stato. Formalmente questo può essere scritto come

$$P(X_{n+1} \leq x_{n+1} | \bigcap_{i=1}^n X_i \leq x_i) = P(X_{n+1} \leq x_{n+1} | X_n \leq x_n)$$

Questa è detta proprietà di Markov, o condizione di assenza di memoria [57].

### 4.5.3 Catena di Markov

Una catena di Markov è un processo di Markov con spazio degli stati discreto, quindi si tratta di un processo stocastico che assume valori in uno spazio discreto e che gode della proprietà di Markov. L'insieme di spazio degli stati può essere finito o infinito (numerabile). Nel primo caso si parla di catena di Markov a stati finiti. Una catena di Markov può essere tempo-continua o tempo-discreta, in base all'insieme di appartenenza della variabile tempo (continuo o discreto). Formalmente, una catena di Markov è una sequenza di variabili aleatorie  $X_1, X_2, \dots$ , che soddisfa la condizione di Markov.

### 4.5.4 Modelli nascosti di Markov

In un modello nascosto di Markov abbiamo due processi stocastici: il processo di interesse inosservabile  $X = (X_s)_{s \in S}$  e il processo osservabile  $Y = (Y_s)_{s \in S}$ . Un modello nascosto di Markov è  $(X, Y, T, E, \pi)$  tale che:

$T : X \times X \rightarrow [0; 1]$  è la probabilità di transizione. In particolare

$$p(X_{t+1} = x_i | X_t = x_j) = T(x_i, x_j)$$

è la probabilità che il processo nascosto passi dallo stato  $x_i$  allo stato  $x_j$ .

$E : X \times Y \rightarrow [0; 1]$  è la probabilità condizionata di  $X$  rispetto ad  $Y$ . Cioè se osserviamo un certo valore  $y$  per il processo osservabile  $Y$  allora  $E$  ci dice qual'è la probabilità che il processo nascosto assuma certi valori.

$\pi : X \rightarrow [0; 1]$  è la probabilità iniziale che il processo nascosto si trovi in un certo stato.

#### 4.5.5 Catena di coppie markoviane

Supponiamo che  $X = (X_1, \dots, X_n)$  e  $Y = (Y_1, \dots, Y_1)$  siano due processi stocastici, nei quali ogni  $X_i$  ha valori in un insieme finito  $\Omega = \{\omega_1, \dots, \omega_n\}$  e ogni  $Y_i$  ha valori reali. Sia  $Z = (Z_1, \dots, Z_n)$  il processo stocastico delle coppie di  $X, Y$ , cioè  $Z_i = (X_i, Y_i)$ . I processi  $X$  e  $Y$  sono una catena di coppie markoviane se  $Z$  è un processo di Markov cioè quando [31] :

$$p(z) = p(z_1)p(z_2|z_1) \cdots p(z_n|z_{n-1})$$

#### 4.5.6 Catena di triple markoviane

Le catene di triple markoviane sono una estensione delle catene di coppie di Markov nelle quali la distribuzione di probabilità di  $(X, Y)$  è una distribuzione di probabilità marginale di una catena di Markov  $(X, U, Y)$ . Anche nelle catene di triple markoviane cerchiamo  $X$  a partire da  $Y$ , il processo  $S$  serve solo come aiuto al calcolo. Sia  $S = N$  l'insieme dei numeri naturali escluso lo zero e siano  $X = (X_n)_{n \in N}$ ,  $Y = (Y_n)_{n \in N}$ ,  $U = (U_n)_{n \in N}$  tre processi, nei quali le variabili  $X_i$  assumono valori in uno spazio campionario  $\Omega = \{\omega_1, \dots, \omega_k\}$ , le variabili  $Y_i$  sono a valori reali e le variabili  $U_i$  assumono valori nell'insieme  $\Lambda = \{\lambda_1, \dots, \lambda_m\}$ . Con lo scopo di semplificare la notazione, poniamo  $T_n = (X_n, U_n, Y_n)$ ,  $Z_n = (X_n, Y_n)$  e  $V_n = (X_n, U_n)$  e chiamiamo  $T, Z$  e  $V$  i processi corrispondenti. Diremo che  $Z$  è una catena di triple di Markov se esiste un processo  $U$  nel quale  $T$  è una catena di Markov. In questo caso allora  $(V, Y)$  è una catena di Coppie di Markov. Possiamo scrivere la distribuzione di probabilità di  $(V_i, Y)$  come

$$p(v_i, y) = \alpha^i(v_i) \beta^i(v_i)$$

con  $\alpha^i(v_i)$  e  $\beta^i(v_i)$  definite rispettivamente come

$$\alpha^i(v_i) = p(y_1, \dots, y_{i-1}, y_i, v_i) \quad \beta^i(v_i) = p(y_{i+1}, \dots, y_n, v_i, y_i)$$

## **Parte II**

### **Stato dell'arte e analisi delle metodologie**



# Capitolo 5

## Stato dell'arte

Nell'esaminare la letteratura ci siamo incentrati sull'analisi del suono respiratorio per poterne determinare le fasi respiratorie e quindi per poter determinare se in un dato istante c'è un respiro o no. Non abbiamo trovato sistemi che si concentrano sul problema dell'apnea notturna o simili cioè che considerino il caso in cui per un determinato lasso di tempo la respirazione sia assente e di conseguenza si è in una situazione di allarme, ma abbiamo trovato sistemi di riconoscimento delle fasi respiratorie svolti tramite diverse metodologie che sono risultati comunque molto utili perché risolvono una parte cruciale del nostro problema. Alcuni di questi sistemi per l'analisi del respiro non usano solamente il suono polmonare ma utilizzano anche quello tracheale ed è stato scelto di analizzare anche questo aspetto perché utile sia per la comprensione degli algoritmi e delle scelte sia per evidenziare le differenze dei due suoni.

### 5.1 Algoritmi di riconoscimento delle fasi e dei cicli respiratori

Ricordiamo che per fase respiratoria intendiamo un evento di inspirazione, di espirazione o di pausa che intercorre tra i due eventi. Con il termine ciclo respiratorio intendiamo una fase di inspirazione seguita da una fase di espirazione seguita da una pausa. Si parla di ciclo respiratorio perché naturalmente è una ripetizione continua delle tre fasi. Una tipica registrazione del suono del torace può essere vista come una sommatoria del suono del cuore, del suono dei polmoni e dei rumori di sottofondo.

### 5.1.1 *Automated respiratory phase and onset detection using only chest sound signal* [36]

In questo articolo viene trattato un algoritmo per la determinazione delle fasi di respirazione analizzando i suoni polmonari. Vengono prese delle registrazioni di suoni polmonari e vengono applicati dei filtri passa banda e non lineari per poter ottenere un suono pulito del respiro. In seguito attraverso l'analisi dello zero crossing (attraversamento della funzione per l'asse dell'ascissa) si riescono a individuare gli onset relativi all'espirazione e all'inspirazione.

**Input** I dati considerati in input sono tre registrazioni per ognuno dei due soggetti sani, di sesso opposto e di anni 25 e 30. In una prima fase i soggetti hanno cercato di mantenere una frequenza di respiro breve, successivamente ai soggetti è stato chiesto di aumentare la frequenza. In questo studio si sono usati solo i suoni registrati sul petto dei soggetti sul quarto interspazio con lo scopo di avere una distinzione ottima tra le varie fasi della respirazione, con una frequenza di campionamento di  $10240Hz$ .

**Algoritmo** Il segnale di input passa attraverso le seguenti fasi:

- Riduzione della sequenza di campionamento a  $1000Hz$  in quanto gli spettri di potenza dei suoni del cuore e dei polmoni normalmente hanno componenti di frequenza inferiori a  $500Hz$ .
- Per l'analisi del flusso, il suono viene pre-elaborato da un filtro passa banda per eliminare il suono inerente al cuore. Tuttavia, a causa della sovrapposizione degli spettri di polmone e toni cardiaci, il filtro passa banda non elimina completamente i toni cardiaci.
- Viene usato un filtro non lineare per eliminare completamente alcuni picchi rimasti nel segnale a causa dei suoni del cuore. La scelta del filtro non lineare è dovuta al fatto che questo permette l'eliminazione dei picchi invece che la loro levigazione.
- L'output appena ottenuto viene sottoposto a un ulteriore filtro passa banda impostato sulla frequenza tipica delle fasi respiratorie. In conclusione si ottiene quindi un suono pulito, dove sono stati scartati i suoni relativi al rumore attraverso il campionamento e grazie ai vari filtri è stato escluso anche il suono cardiaco.

Nell'immagine 5.1 viene illustrato il risultato di questi passaggi, dove la linea blu rappresenta il suono filtrato, mentre la linea rossa rappresenta la frequenza tipica respiratoria usata come parametro nell'ultimo filtro usato.

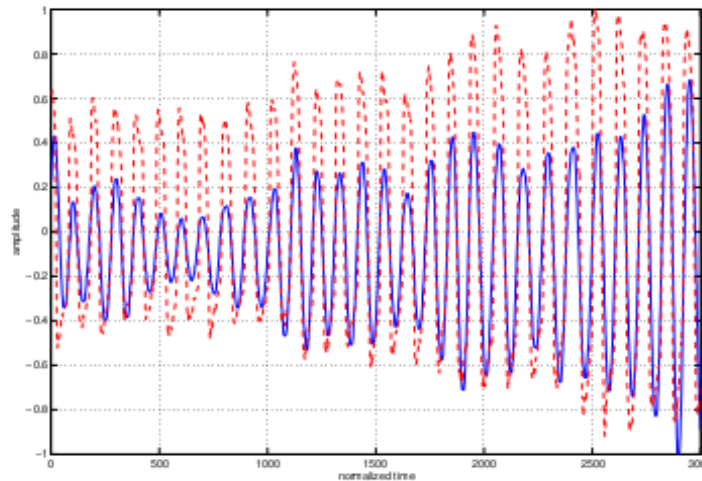


Figura 5.1: Suono del respiro dopo applicazione di un filtro passa banda impostato sulla frequenza tipica delle fasi respiratorie

- Si effettua un'analisi dello zero crossing usato per stimare gli onset, infatti si può notare che le rispettive fasi di inspirazione ed espirazione avvengono nei passaggi per lo zero: dove il segno del segnale cambia da positivo a negativo inizia una fase di espirazione mentre un cambiamento di segno da negativo a positivo corrisponde a un inizio di inspirazione.
- Nell'analisi finale si scelgono i picchi attraverso un thresholding adattativo. Viene applicato un filtro passa banda incentrato sull'approssimazione del tasso di respirazione dell'energia del segnale: il suono viene suddiviso in finestre temporali così che i punti di rapida diminuzione o aumento di energia possano essere meglio analizzati e usati per la raffinazione della stima precedente.

**Conclusioni** Il ritardo medio di questo algoritmo nel riconoscere gli onset della fase respiratoria non supera i  $100ms$ , ed è un intervallo di tempo assolutamente accettabile ai fini della diagnosi in tempo reale della sindrome da apnea del sonno. Tuttavia l'algoritmo deve essere modificato per trattare variabili pattern respiratori e nel caso in cui i soggetti dormano, alcuni parametri devono essere modificati. Lo studio non riporta nessun dato sulla complessità computazionale e sull'implementazione dell'algoritmo.

### 5.1.2 *Breath Analysis of Respiratory Flow using Tracheal Sounds* [20]

In questo studio gli autori studiano le differenze che ci sono tra la fase inspiratoria e la fase espiratoria in due quantità relative a un segnale tracheale filtrato con un filtro passa banda. Queste due quantità sono la media e la varianza logaritmica dell'energia. Lo

studio inoltre usa anche uno spirometro per misurare il flusso. Il flusso viene diviso in base al valore assoluto in: basso, medio, alto e molto alto. Questo algoritmo quindi non ricava il flusso a partire dal suono ma è utile per comprendere e sviluppare algoritmi del tipo che ci interessa.

**Input** I dati presi in input sono delle registrazioni di suoni tracheali registrati su nove soggetti sani e non fumatori i quali non hanno mai avuto gravi malattie respiratorie. Inoltre lo studio aveva a disposizione anche il flusso d'aria registrato attraverso uno spirometro con pneumotacografo.

**Algoritmo** L'algoritmo ha due flussi di esecuzione indipendenti, il primo è il seguente:

- Filtro passa alto con frequenza di taglio di  $70Hz$  per rimuovere il rumore a bassa frequenza.
- Nelle fasi seguenti l'algoritmo considera solo le porzioni del suono registrate quando il segnale del flusso era al di sotto del 20% del flusso medio o al di sopra del 20% di esso. Perché in queste condizioni il suono tracheale si può considerare stazionario.
- Lo spettro di potenza dei suoni della trachea è stato calcolato in una finestra di  $50ms$  (512 campioni) con il 75% di sovrapposizione tra finestre successive. Per ogni fase respiratoria durante la quale c'era un diverso flusso d'aria, è stata calcolata la media della potenza dei suoni tracheali in decibel entro sei predefinite bande di frequenza: da  $70$  a  $300Hz$ , da  $300$  a  $450Hz$ , da  $450$  a  $600Hz$ , da  $600$  a  $800Hz$ , da  $800$  a  $1000Hz$  e da  $1000$  a  $1200Hz$ .
- Dato che le intensità dei suoni respiratori variano da soggetto a soggetto, per ogni soggetto i valori calcolati in precedenza sono stati normalizzati rispetto al valore massimo.
- Si è poi calcolata la media dei valori normalizzati tra soggetti diversi per ogni fase respiratoria, inoltre sono state calcolate la media e l'errore standard per diversi tassi di flusso e intervalli di frequenza.

Mentre il secondo flusso di esecuzione è:

- Il segnale dei suoni della trachea sono stati filtrati attraverso un filtro passa alto nelle stesse frequenze menzionate in precedenza.



- Il segnale filtrato è stato in seguito segmentato in finestre di dimensione  $50ms$  (512 campioni) con il 75% di sovrapposizione tra finestre successive usando una finestra di Hanning.
- Si calcola il logaritmo della varianza dei segmenti precedenti.
- In ciascuna finestra il valore precedente viene normalizzato rispetto al valore massimo per ridurre le interferenze dei suoni del cuore.
- In seguito viene calcolata la media all'interno delle diverse bande di frequenza e dei diversi valori del flusso d'aria.

**Conclusioni** Da una analisi dello spettrogramma dei suoni tracheali, si può vedere che l'intensità del suono tracheale aumenta con l'aumentare del valore assoluto del flusso. Inoltre anche la varianza logaritmica normalizzata e la media di potenza normalizzata seguono i cambiamenti nel valore assoluto del flusso. Nello spettrogramma, nella varianza logaritmica normalizzata e nella media normalizzata della potenza sono evidenti le transizioni di fase respiratoria. Quando il flusso era medio o alto, si ha una maggiore differenza di media dell'energia normalizzata tra la fase inspiratoria ed espiratoria nella banda di frequenze dai 300 ai  $450Hz$ . Inoltre questa banda di frequenza ottiene la seconda maggior differenza di media dell'energia normalizzata tra la fase inspiratoria e la fase espiratoria quando il flusso è basso o molto alto. Quindi questo intervallo di frequenza è stato scelto come ottimale per esaminare i cambiamenti nella media della potenza rispettivamente alle fasi respiratorie.

### 5.1.3 *Acoustical respiratory signal analysis and phase detection* [21]

Questo articolo propone un approccio di modellazione statistica per il riconoscimento delle fasi respiratorie.

**Input** L'input usato per i test è preso da due segnali respiratori di due pazienti nell'intervallo di frequenza dai 250 ai  $312.5Hz$ . Il primo dura 16s e contiene circa tre cicli di respirazione, il secondo dura 8 secondi e contiene anch'esso tre cicli.

**Algoritmo** Prima di tutto l'analisi del suono viene fatta nel dominio dei pacchetti wavelet per incrementare l'accuratezza della determinazione dei suoni. Il sistema di riconoscimento è implementato attraverso una segmentazione del segnale in tre parti alle quali viene associata una etichetta che può essere: inspirazione, espirazione e transizione. Viene adottata una rete Bayesiana, usando una versione con vincoli di una catena di Markov

a triple. Questo modello sfrutta un ciclo respiratorio a priori, con lo scopo di guidare l'algoritmo verso un accurato riconoscimento delle fasi della respirazione.

**Conclusioni** I risultati sono buoni ma è presente qualche errore di classificazione. Questo a causa del rumore ambientale nella prima registrazione e a causa della forte variazione di intensità tra le fasi espiratorie nella seconda registrazione.

#### **5.1.4 *Computerized acoustical respiratory phase detection without air-flow measurement* [26]**

Questo studio sviluppa un metodo per riconoscere le fasi respiratorie a partire dai suoni prodotti dall'apparato respiratorio. Viene implementato anche un programma MATLAB.

**Input** Sono stati studiati 21 soggetti di età dai 4 ai 51 anni. I soggetti sono stati divisi in gruppi di età: 7 bambini e 4 bambine con una età media di 10 anni e 4 uomini e 6 donne con una età media di 32 anni. Tutti i soggetti godevano di buona salute e non avevano infezioni delle vie respiratorie nelle quattro settimane precedenti alla registrazione. Per la registrazione dei suoni respiratori sono stati usati sei accelerometri. I dispositivi di registrazione erano attaccati sulla trachea e in altri punti del petto. Viene anche registrato in modo simultaneo, il flusso d'aria attraverso uno pneumotacografo con un trasduttore di differenza di pressione. L'algoritmo prende in input solo i suoni della registrazione mentre il flusso d'aria registrato dallo pneumotacografo serve esclusivamente per valutare la qualità dei risultati dell'algoritmo.

#### **Algoritmo**

- Il segnale è stato amplificato, filtrato attraverso un filtro passa banda con banda dai 50 ai 200Hz e digitalizzato a una frequenza di campionamento pari a 10240Hz.
- Una finestra di Hanning è stata usata per segmentare il segnale sonoro in finestre di 2048 campioni con il 50% di sovrapposizione tra segmenti successivi.
- Lo spettro di potenza di ogni segmento è stato calcolato con una trasformata veloce di Fourier. Le bande di frequenza usate per calcolare la potenza media di ogni segmento erano: dai 150 ai 300Hz, dai 300 ai 450Hz, dai 450 ai 600Hz e dai 600 ai 1200Hz. L'algoritmo in modalità di esecuzione semiautomatica offre all'utente la possibilità di visualizzare i dati e di identificare eventuali suoni accidentali da rimuovere.

- Per ogni segmento, lungo  $200ms$ , viene calcolata la media della potenza del segnale registrato sulla parete del petto nella posizione con la quale si ottiene la maggiore differenza di suono tra inspirazione ed espirazione. Questo è risultato in un segnale nel quale i picchi corrispondono al massimo flusso d'aria durante l'inspirazione.
- È stata usata una finestra mobile per riconoscere i picchi inspiratori per il segnale completo. Per venire in contro alla variabilità nella frequenza respiratoria, la lunghezza della finestra è stata scelta in modo tale da approssimare la durata di un ciclo di respirazione. Dato che il suono polmonare è più forte durante l'inspirazione che durante l'espirazione, l'algoritmo usa i picchi del suono polmonare per determinare i picchi di inspirazione. Invece i suoni tracheali sono forti sia durante l'inspirazione che durante l'espirazione e quindi li usa per determinare gli onset delle fasi respiratorie.
- Infine l'algoritmo classifica come inspirazione quella regione temporale che è compresa tra due onset tracheali e che contiene un picco polmonare. Tutto il resto viene classificato come espirazione.

**Conclusioni** Il software ha ottenuto una accuratezza massima nella stima delle fasi respiratorie senza l'uso dei dati sulla misurazione diretta del flusso d'aria. La combinazione dei suoni polmonari e quelli tracheali ha facilitato il riconoscimento delle fasi respiratorie. Sfruttando la differenza di suono polmonare tra inspirazione ed espirazione l'algoritmo è riuscito a distinguere le varie fasi respiratorie mettendo in relazione i picchi di suono tracheale con quelli del suono polmonare.

### **5.1.5 *Respiratory onset detection using variance fractal dimension* [22]**

Anche [22] si occupa del problema di sviluppare un metodo acustico non invasivo per riconoscere le fasi respiratorie senza una misurazione diretta del flusso d'aria e lo fa usando la dimensione frattale. Questa è una misura della complessità di un segnale. Una proprietà della dimensione frattale è che è indipendente dalla potenza del segnale. Questo sistema si basa sull'assunto che durante la transizione di fase respiratoria, il segnale sonoro ha un comportamento caotico a causa del cambiamento di momento nel flusso d'aria mentre questo cambia di direzione. Quindi si ipotizza che la varianza della dimensione frattale del suono respiratorio abbia un picco negli onset dei cicli respiratori.

**Input** Lo stesso input utilizzato dallo studio [26], il quale è già stato descritto in precedenza.

**Algoritmo** Per trovare gli onset dei cicli respiratori, viene calcolata la varianza della dimensione frattale usando un segmento di 128 campioni pari a  $12.5ms$  con il 50% di sovrapposizione tra segmenti adiacenti. In seguito viene usata una finestra di lunghezza pari alla durata approssimativa della metà di una singola fase respiratoria, cioè  $0.7s$ , con lo scopo di riconoscere i picchi nella varianza della dimensione frattale.

**Conclusioni** Da un confronto tra il riconoscimento degli onset con il reale flusso d'aria i risultati mostrano che l'intervallo di errore va dai  $31ms$  ai  $49ms$ . L'aspetto positivo di questo algoritmo è che fa una analisi esclusivamente nel dominio del tempo del segnale.

### 5.1.6 *Automated respiratory phase detection by acoustical means* [7]

Lo studio [7] si concentra sull'automazione del processo di riconoscimento per via acustica delle fasi della respirazione senza l'ausilio di misure del flusso. Vengono usati due tipi di suoni: quello tracheale e quello del torace, i primi vengono usati per determinare gli onset mentre i secondi per riconoscere le fasi respiratorie. Viene implementato un programma in C++ dotato di una interfaccia grafica in grado di lavorare in moto semi automatico o completamente automatizzato.

**Input** L'input è costituito da un insieme di 17 suoni registrati sulla trachea e sul torace (sul secondo interspazio sinistro medioclavicolare e sul terzo interspazio destro medioclavicolare) di 11 soggetti sani con età dai 4 ai 35 anni. Inoltre erano disponibili anche i dati del flusso d'aria misurati attraverso uno pneumotacografo.

#### **Algoritmo**

- Il segnale del suono della respirazione viene segmentato in segmenti di lunghezza  $100ms$ , ognuno con una sovrapposizione del 50% tra segmenti adiacenti.
- Per ogni segmento viene calcolato lo spettro di potenza usando una trasformata veloce di Fourier.
- I segnali del torace vengono filtrati lasciando una banda di frequenza dai  $150Hz$  ai  $300Hz$  mentre i segnali della trachea vengono filtrati lasciando una banda di frequenza dai  $150Hz$  ai  $600Hz$ .
- Viene presa la potenza media dei segnali.
- Per riconoscere i picchi di inspirazione dai segnali del petto, inizialmente, viene calcolata per ogni campione la pendenza della curva del segnale, prendendo la differenza tra punti adiacenti. I campioni che hanno una differenza positiva con il

campione precedente e una differenza negativa con il campione successivo sono i possibili picchi.

- Per affinare la ricerca dei picchi inspiratori, viene usata una piccola finestra mobile. Dato che il valore del flusso può variare molto da respiro a respiro, la finestra è stata scelta in modo da coprire un ciclo completo di respirazione nel segnale (approssimativamente  $2s$  quindi 20 campioni). Una finestra di questo tipo è applicata ai picchi determinati al punto precedente. Vengono determinati i punti massimi della finestra che hanno almeno una deviazione standard al di sopra della media. Ci si aspetta che i picchi di inspirazione siano molto più alti dei picchi di espirazione per lo spettro di potenza dei suoni registrati sul petto. Poiché i possibili picchi possono consistere sia in picchi di inspirazione che in picchi di espirazione, bisogna usare una soglia per eliminare i picchi di espirazione. L'algoritmo di riconoscimento dei picchi è progettato per essere usato sia in modalità automatica che in modalità semiautomatica. Nella modalità semiautomatica, gli utenti possono aggiungere i picchi mancanti o eliminare i picchi riconosciuti dopo aver dato un'occhiata alla potenza media dello spettro dei segnali.
- Per trovare gli onset del respiro dai segnali tracheali si calcola la pendenza di ogni campione: i campioni che hanno una differenza negativa rispetto ai campioni precedenti e una differenza positiva rispetto ai campioni successivi vengono selezionati come potenziali onset. Dato che in media la durata delle fasi respiratorie è approssimativamente  $1s$ , la distanza tra un picco e l'onset più vicino è di circa  $500ms$  o 5 campioni. Viene considerata quindi troppo vicina una differenza di 2 campioni ( $200ms$ ), dunque i potenziali onset riconosciuti che sono considerati troppo vicini a uno dei picchi precedenti vengono esclusi dalla selezione.
- Per calcolare il valore medio e la deviazione standard di ogni campione, si usa una finestra mobile lunga 10 campioni (approssimativamente metà della durata di un respiro). I potenziali onset con deviazione standard di almeno 0.5 più piccola della media locale vengono scelti per una successiva analisi.
- Nel processo di riconoscimento dei picchi tramite i risultati ottenuti viene stimata la distanza media tra due picchi inspiratori. La lunghezza media di una fase respiratoria è calcolata come la metà della media della distanza tra i picchi. Questa informazione è usata per raggruppare i potenziali onset in cluster. Si classifica come onset del respiro solo il punto minimo all'interno di ogni cluster.
- Per massimizzare l'accuratezza del programma di riconoscimento della fase respiratoria viene fatta un'ulteriore ottimizzazione aggiungendo un algoritmo di auto-

correzione. Questo algoritmo confronta gli onset trovati con i picchi stimati per assicurarsi che esistano solo due onset tra due picchi inspiratori. Gli onset che non vengono riconosciuti durante il processo di selezione precedente vengono aggiunti mentre gli onset extra che esistono tra due picchi vengono rimossi. Con i picchi stimati e gli onset si possono prevedere le fasi respiratorie o la direzione del flusso d'aria. Gli intervalli tra due onset con un picco sono inspiratori mentre gli intervalli senza picchi tra gli onset sono espiratori.

**Conclusioni** Per calcolare le prestazioni complessive del sistema sono stati valutati tre parametri:

- Percentuale delle fasi respiratorie riconosciute in modo corretto dall'algoritmo di riconoscimento dei picchi.
- Percentuale degli onset riconosciuti in modo corretto dall'algoritmo di riconoscimento degli onset.
- Media delle differenze in millisecondi tra gli onset riconosciuti e quelli trovati direttamente dalla misurazione del flusso.

Nella modalità semiautomatica i picchi di inspirazione vengono riconosciuti con una accuratezza del 100%. In modalità completamente automatica l'accuratezza media è del 93% con una deviazione standard del 7%. Confrontato con il flusso misurato direttamente, in media, c'è solo un ritardo di  $118ms$  di riconoscimento tra gli onset rilevati e quelli reali. Dato che è inevitabile un minimo flusso critico all'interno della trachea per generare un suono udibile dallo stetoscopio, ci si aspetta una differenza tra gli onset reali dei cicli della respirazione ottenuti attraverso uno pneumotacografo e gli onset stimati dal sistema. Studi precedenti mostrano che questa differenza è di circa  $40ms$  [26]. Questo sistema stima gli onset dei cicli respiratori e delle fasi respiratorie e la direzione del flusso d'aria ma non tenta di stimare anche il valore assoluto del flusso.

### **5.1.7 A Software Toolkit for Acoustic Respiratory Analysis [2]**

L'obiettivo principale di questo studio è la realizzazione di un software per il riconoscimento delle fasi respiratorie e per la classificazione dei suoni respiratori.

**Input** Usa i suoni tracheali di cinque soggetti sani, registrati con uno stetoscopio elettronico a una frequenza di campionamento di  $22050Hz$ .

## Algoritmo

- La fase di estrazione dei suoni polmonari rimuove la componente del segnale le cui frequenze si trovano al di fuori della banda di frequenza che va dai 100 ai 2500Hz, dato che i suoni respiratori tracheali di solito non cadono al di fuori di questa banda. Per farlo viene usato un filtro passa banda appropriato. Questo filtro usa una finestra di Blackman di dimensione pari a un dodicesimo della frequenza di campionamento del segnale di input. Questo filtro è importante soprattutto per rimuovere i suoni ad alta frequenza che sono di origine cardiocircolatoria.
- Una fase successiva prende il valore assoluto del segnale.
- Per ridurre la complessità computazionale viene applicato al segnale un sottocampionamento di fattore 60 applicando una finestra mobile di 60 campioni che prende quindi un campione ogni 60. Lo scopo della fase di trattamento del segnale è quello di trovare la forma generale del segnale. Il sottocampionamento non degrada la forma di questa a patto che la nuova frequenza di campionamento sia almeno il doppio della frequenza di respirazione che si può evincere tipicamente dai suoni respiratori tracheali. Se assumiamo che la durata tipica di una fase respiratoria è di un secondo allora l'envelope (2.4) del segnale sarà di 1Hz. Questo significa che la nuova frequenza di campionamento deve essere di almeno 2Hz. In questo caso la nuova frequenza di campionamento è di circa 367Hz.
- Il segnale attraversa una fase di riduzione del rumore che usa un filtro mediano per ridurre il valore assoluto dei picchi di durata più breve. Il segnale viene prima partizionato in segmenti contigui di 0.5s e per ogni segmento viene calcolato il valore mediano. Si confrontano i campioni di ogni segmento con un valore di soglia pari a cinque volte il valore mediano del segmento a cui appartengono, se il campione supera questa soglia viene azzerato perché considerato rumore. Questo procedimento viene ripetuto due volte.
- Il segnale viene nuovamente sottocampionato di un fattore due e passa attraverso un filtro passa basso con una finestra di Hamming di 50ms. Lo scopo di questo filtro è di ridurre i suoni che hanno una varianza alta e una frequenza alta.
- Viene applicato un filtro con una impulse response di lunghezza dipendente dalle caratteristiche del segnale di input. La lunghezza  $w$  della finestra usata dal filtro approssima la media della durata di una fase respiratoria nel segnale di input. Per determinare la lunghezza del filtro, cominciamo con una fase di detrend: si sottrae la media globale del segnale di input da ogni campione in modo tale che la media dell'output sia zero.

- Per approssimare la media della durata di una fase respiratoria ci si basa sull'assunto che questa sia proporzionale alla ampiezza dei picchi. La fase di calcolo della durata media di una fase respiratoria, approssima questa quantità trovando il picco più ampio che interseca la linea di ampiezza zero. Supponiamo che il segmento originato da questa intersezione abbia lunghezza  $L$ . Allora la dimensione  $w$  della finestra è una funzione di  $L$  ed è scelta in modo empirico e arbitrario nel modo seguente:  $w = f(L) = (w/2) + 2$  dove  $(w/2)$  viene arrotondato per eccesso a un numero intero. Solo a questo punto possiamo applicare il filtro passa basso con una finestra di Hamming di lunghezza  $w$ .
- Usiamo le informazioni nel dominio del tempo per classificare il suono di input in respiratorio o non respiratorio. Questa fase usa il segnale della fase precedente e la dimensione  $w$  della finestra che approssima la durata media di una fase respiratoria.
- Il segnale viene diviso in segmenti contigui di lunghezza  $3w$  e viene calcolata l'energia totale  $E$  di un segmento sommando i valori di energia dei segmenti appena raggruppati.
- L'energia  $E$  viene confrontata con una soglia uguale all'energia di base del segmento. Definiamo l'ampiezza di base come la media delle ampiezze nell'area del segnale nella quale non c'è respirazione e definiamo l'energia di base  $E_{baseline}$  come la lunghezza del segmento moltiplicata per l'ampiezza di base. Viene usato un valore fisso determinato empiricamente per l'ampiezza di base. Se l'energia  $E$  supera l'energia di base  $E_{baseline}$  allora il segmento viene classificato come respiro altrimenti come pausa. Viene prodotta l'etichetta 1 se c'è respiro altrimenti viene prodotta l'etichetta 0.
- Le etichette prodotte nella fase precedente vengono raggruppate se uguali e contigue.
- Un respiro è un gruppo di etichette 1 seguito da un gruppo di etichette 0.

**Conclusioni** Il sistema appena analizzato non si concentra solo sul riconoscimento delle fasi respiratorie. Lo scopo finale di questo algoritmo consiste nel riconoscimento dei suoni anormali respiratori come i crackles e i sibili per poterne studiare le differenze rispetto a un suono respiratorio normale. L'algoritmo che abbiamo appena spiegato quindi consiste solo nella fase iniziale del sistema, ed è la fase di nostro interesse. È stato definito un livello di accuratezza pari al 97% per il riconoscimento dei suoni anormali respiratori.



## 5.2 Classificazione degli algoritmi

In questa sezione ci occupiamo della classificazione degli algoritmi appena visti. La prima distinzione che si può fare è quella riguardante la metodologia usata nella soluzione del problema. In particolare si nota che gli articoli [36] [26] [7] [2] [22] usano tutti una tecnica di beat detection per analizzare il respiro e ricavare gli onset della respirazioni, applicano dei filtri passa banda per l'eliminazione del rumore e del battito cardiaco e usando un threshold adattivo per rilevare gli onset.

In particolare abbiamo che i sistemi [26] [7] sono molto simili tra loro perché entrambi dopo aver applicato dei filtri passa banda suddividono l'onda in segmenti e ne calcolano la potenza media infine tramite una finestra mobile con una durata pari a un ciclo respiratorio riconoscono gli onset della respirazione. Le differenze tra questi due sistemi sono per l'appunto minime come per esempio la scelta della dimensione dei segmenti di analisi: Il primo utilizza un segmento di  $200ms$ , mentre il secondo di  $100ms$ , più altri piccoli dettagli.

L'articolo [36] invece usa una tecnica di riconoscimento differente: oltre a un filtraggio passa banda usa anche un filtro non lineare per l'eliminazione totale del cuore e un filtro impostato sulle frequenze respiratorie. Per la stima degli onset invece viene usata l'analisi dello zero-crossing in quanto le rispettive fasi di inspirazione ed espirazione avvengono ai passaggi per lo zero.

L'articolo [2] dopo particolari filtri e downsampling usa come il precedente la tecnica di zero-crossing, ma la utilizza in modo differente: viene usata per determinare il ciclo di fase respiratoria, supponendo che in un ciclo di respirazione sono presenti un numero prefissato di passaggi per lo zero dell'onda. Questa tecnica viene usata quindi per determinare la finestra di analisi alla quale viene poi calcolata l'energia totale e confrontata con una soglia fissa determinata empiricamente; se l'energia totale del segmento supera la soglia fissa il segmento viene classificato come respiro.

Infine l'articolo [22] risolve il problema della determinazione degli onset usando la dimensione frattale. Come gli altri sistemi basati sull'onset detection suddivide il suono in segmenti e sfrutta una finestra di lunghezza pari a mezzo respiro. Viene scelta questa tecnica in quanto si ipotizza che la varianza della dimensione frattale del suono respiratorio ha un picco negli onset dei cicli respiratori. Un'altra differenza di questo sistema consiste nel fatto che l'analisi è esclusivamente nel dominio del tempo, a differenza di tutti gli algoritmi di onset detection.

Oltre alla tecnica di onset detection abbiamo analizzato un sistema di natura probabilistica che sfrutta le catene di Markov. L'articolo [21] associa a ogni fase respiratoria, una variabile e attraverso la tripla delle catene Markoviane, riesce ad individuare le fasi

respiratorie nell'onda sonora sfruttando il modello di un ciclo respiratorio conosciuto a priori. I risultati di questo algoritmo sono molto influenzabili dal rumore esterno e da rumori inaspettati emessi dal soggetto, in conclusione non risulta un algoritmo ottimale a meno che il suono prima di essere analizzato venga sottoposto a un miglior filtraggio.

Viene analizzato anche l'articolo [20] che può risultare utile per un ipotetico sviluppo del problema in quanto calcola l'energia media e la varianza logaritmica dei suoni tracheali tra le fasi respiratorie e riesce a riconoscere le fasi di respirazione in base ai cambiamenti di questi valori.

## 5.3 Conclusioni

Questi sistemi di analisi risultano tutti molto utili per l'analisi delle fasi respiratorie. Nessuno di questi però contiene un'analisi nel caso in cui il respiro sia assente. Solo [2] classifica i segmenti come respiro o non respiro, però la sua analisi non viene fatta a runtime, cioè per determinare le soglie per la classificazione ha bisogno di analizzare l'intero file audio contenente il suono respiratorio. In più utilizza solo suoni tracheali, mentre nella nostra analisi del problema è stato scelto di utilizzare i suoni polmonari. Nonostante le differenze dell'input e l'analisi a posteriori del suono questo algoritmo è risultato il più utile per la risoluzione del nostro problema.

Un'importante problematica che non viene affrontata in questi articoli sono i rumori sonori provenienti dal soggetto come per esempio la deglutizione nel caso in cui si stiano analizzando suoni tracheali ed eventuali suoni dovuti a una respirazione anormale (crackles, sibili, rantoli ecc..) che in alcuni casi possono rappresentare un disturbo durante la classificazione. Inoltre essendo tutte registrazioni fatte mentre il soggetto era sveglio non si va a considerare la respirazione notturna che è solitamente più lenta e varia in base alle fasi del sonno.

Tutti i sistemi analizzati hanno anche la relativa implementazione software, usando per esempio MATLAB e C++, riportando anche eventuali risultati nei paper. Spesso non si accenna alle tecniche di implementazione utilizzate e soprattutto nessuno rende disponibile il proprio operato per poter confrontare i risultati o effettuare delle prove con dei suoni respiratori.

## Capitolo 6

### Analisi delle Metodologie

In generale un sistema implementato per il riconoscimento del respiro tramite segnali fisiologici continui deve essere sensibile ai cambiamenti caratteristici dei segnali che si stanno analizzando. Questi cambiamenti dipendono dalla frequenza della respirazione, dal suo volume, dal suo flusso e da altre particolari patologie o eventi che possono essere presenti. Ci si aspetta che un buon algoritmo consideri queste caratteristiche e i conseguenti cambiamenti che possono presentarsi nel tempo durante una analisi della respirazione. Di conseguenza un algoritmo ottimale deve basarsi su delle quantità statistiche del segnale o su qualche forma di apprendimento automatico rispettando un qualche principio di località, in quanto alcune proprietà respiratorie cambiano nel lungo termine.

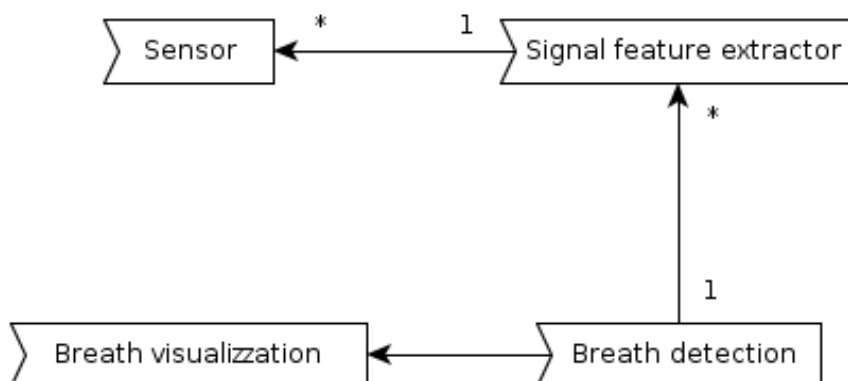


Figura 6.1: Diagramma di implementazione di un sistema per l'analisi del respiro

## 6.1 Metodo elementare

Si può essere tentati a risolvere il problema con un algoritmo elementare come per esempio l'algoritmo 3. Anche nell'ipotesi che non ci sia alcun rumore ambientale non si ha la certezza che tale algoritmo funzioni a causa della presenza di altri rumori fisiologici che potrebbero essere rilevati dallo stetoscopio, come per esempio i suoni cardiovascolari o i suoni gastrointestinali.

L'algoritmo per di più è sottospecificato in quanto ci sono vari gradi di libertà su alcuni parametri che non si è in grado di scegliere, come la scelta della soglia, della dimensione della finestra e dei dati disponibili riguardo ai suoni registrati da uno stetoscopio.

Si ritiene quindi che almeno un primo livello di trattamento del segnale sia necessario per risolvere questo problema, inoltre un'implementazione come quella appena descritta la si può suggerire per una possibile realizzazione a un livello più basso ad esempio direttamente nel firmware dello stetoscopio.

<pre><b>input</b> : A block of sound samples  1 <b>foreach</b> <i>window w in the input block</i> <b>do</b> 2   <i>instantSoundEnergy</i> <math>\leftarrow</math> <i>w.getInstantSoundEnergy()</i> 3   <b>if</b> <i>instantSoundEnergy</i> &gt; <i>threshold</i> <b>then</b> 4       <i>is a breath</i>; 5   <b>else</b> 6       <i>is not a breath</i>;</pre>
--

**Algorithm 3:** Naive breath detection

## 6.2 Beat detection

Analizzando il suono respiratorio si può facilmente notare che il ciclo respiratorio in quanto tale è ciclico e lo si può considerare come un evento periodico come il ritmo musicale per sua stessa definizione. Di conseguenza si possono utilizzare gli algoritmi di beat detection nati per lo studio del ritmo musicale. Osservando i vari algoritmi di beat detection infatti si può notare come possano facilmente essere applicati ai suoni respiratori.

### 6.2.1 Scelta dell'algoritmo

Come è già stato spiegato nel capitolo 4.1 per la determinazione del ritmo bisogna prima trovare gli onset, che rappresentano l'inizio di un suono o di una nota.

Secondo [30] ci sono varie proprietà da considerare per scegliere un algoritmo di onset detection. Ad esempio: la complessità dell'algoritmo, le caratteristiche della piattaforma sulla quale ci si aspetta che l'algoritmo venga usato, la presenza o meno di vincoli sul tempo di esecuzione e altri. È molto importante anche il dominio dell'input: se il segnale ha dei beat molto marcati e presenta relativamente pochi rumori allora un metodo che analizza il dominio del tempo risulta adeguato; se invece il segnale da analizzare è più complesso, dove per esempio sono presenti diversi suoni, conviene usare un metodo basato su informazioni di fase nel dominio delle frequenze; se è necessaria una localizzazione precisa nel tempo e nelle frequenze allora si possono usare metodi basati sulle wavelet; se la complessità computazionale non è un problema ed è presente un insieme adatto di segnali di allenamento allora si possono usare metodi basati su apprendimento automatico e informazioni statistiche (reti neurali, support vector machine, modelli nascosti di Markov).

Si ricorda che gli algoritmi di onset detection sono progettati per funzionare su brani musicali. Questi possono essere un insieme molto complesso di voci musicali. Nell'analisi del suono respiratorio le ipotesi sull'input sono più semplici poiché si può paragonare il suono prodotto dal respiro a un brano musicale relativamente semplice. Quindi si è concluso che è sufficiente una analisi del segnale nel dominio del tempo.

## **6.3 Pattern recognition e apprendimento automatico**

Un'altra possibile soluzione attinge al campo del riconoscimento vocale. In particolare alcune tecniche di riconoscimento vocale usano dei classificatori che hanno come mattoni di base i fonemi. Il suono della respirazione in un certo senso può essere pensato come un linguaggio parlato nel quale ci sono solo due tipi di fonemi: l'inspirazione e l'espirazione. Sia nel linguaggio parlato che nella respirazione è anche importante il riconoscimento del silenzio.

### **6.3.1 Reti neurali**

Una soluzione di questo tipo si può implementare attraverso le reti neurali artificiali. In generale si può procedere nel modo seguente:

- Prima di tutto bisogna disporre di un database di registrazioni di suoni respiratori.
- I suoni possono attraversare una fase di filtraggio nella quale si rimuove il rumore, si riduce eventualmente la frequenza di campionamento e si cerca di eliminare i suoni cardiovascolari.

- Questi suoni si dividono in segmenti di due tipi:
  - suoni che contengono una respirazione, cioè i suoni che contengono il suono prodotto dall’espiazione o dall’inspirazione.
  - suoni che contengono pause respiratorie. Per pause respiratorie in questo caso si intende una pausa in generale e quindi non necessariamente una apnea.
- Per ogni segmento si calcola una sequenza di  $S$  proprietà statistiche interessanti ad esempio: la media dei valori assoluti dei coefficienti di Fourier in determinate bande di frequenza, la potenza media dei coefficienti wavelet in determinate bande di frequenza, la deviazione standard dei coefficienti in determinate bande di frequenza e il rapporto tra la media dei valori assoluti in bande di frequenza adiacenti.
- Si calcola inoltre la quantità  $l$  che è la media della durata dei segmenti che contengono suoni respiratori.
- A questo punto si usano i valori calcolati in precedenza per allenare una rete neurale. Una rete neurale è un meccanismo per approssimare una funzione quindi per allenare una rete neurale bisogna fornire ad essa delle coppie (*input*, *output*). In questo caso possiamo usare due neuroni di output e un numero di neuroni di input pari ad  $n$ , dove  $n$  è il numero di proprietà dei segmenti di suono calcolate in precedenza. Possiamo costruire un insieme di allenamento nel modo seguente. Se  $S$  è una sequenza relativa a un segmento di respiro allora inseriamo nell’insieme  $(S, (1, 0))$  mentre se  $S$  è una sequenza relativa a un segmento di pausa inseriamo nell’insieme  $(S, (0, 1))$ .
- Dopo che la rete è stata allenata, la si può usare per classificare un segnale di input. Come prima si divide il segnale in segmenti di dimensione pari a  $l$ .
- Si dà il segmento in pasto alla rete neurale che produce l’output:  $(s, p)$ .
- Se  $s$  è maggiore di  $p$  allora il segmento è un suono, altrimenti è una pausa.

## 6.4 Clustering

Guardando il grafico della forma d’onda del segnale della respirazione, in assenza di forte rumore, nella giusta scala temporale ed entro certi limiti di quantità temporale del segnale, siamo in grado di dividere il grafico in parti regolari e cicliche e di classificarle in: espiazione, inspirazione, pausa. Questo approccio metodologico risulta particolarmente chiaro osservando il grafico in figura 6.2.

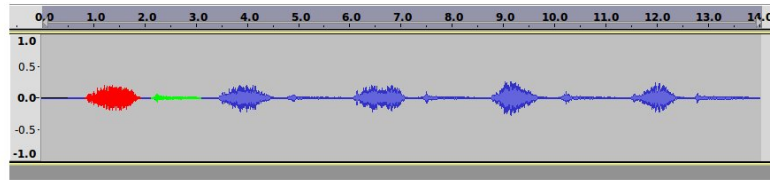


Figura 6.2: Forma d'onda del suono respiratorio

Tale processo intuitivo si può inquadrare nel problema più generale del clustering. Possiamo procedere nel modo seguente:

- Dividere l'input in blocchi di una lunghezza abbastanza grande da essere sicuri che contenga un ciclo respiratorio completo come per esempio di 20 secondi.
- Trovare un clustering adatto ai dati di input.
- Un ciclo respiratorio si può definire con due cluster: uno contenente inspirazione ed espirazione e il secondo contenente la pausa, oppure si può dividere ulteriormente il primo cluster separandolo in due cluster: uno per l'inspirazione e uno per l'espirazione. Se durante l'analisi i cluster si alternano allora il soggetto sta respirando.

Rimane il problema sulla scelta dell'algoritmo di clustering adatto alla nostra situazione. Nel nostro caso siamo di fronte a un problema di clustering nel quale lo spazio del segnale di input è bidimensionale. Una dimensione è il tempo discreto e l'altra dimensione è il valore dei campioni.





# Capitolo 7

## Meccanismi di valutazione

Per poter realizzare un sistema ottimale per il riconoscimento dell'apnea notturna tramite monitoraggio attraverso uno stetoscopio elettronico bisogna tener conto di alcuni requisiti essenziali del programma. Oltre a una verifica del software per assicurarsi che sia corretto e affidabile e altre verifiche come la complessità computazionale e l'utilizzo del processore sono molto importanti dei test valutativi per verificare l'effettiva efficienza del sistema.

### 7.1 Classificazione dello spazio dell'input

Il tipo dell'input è l'insieme di tutti i possibili segnali audio di lunghezza non superiore alle 12 ore. Alcuni file audio utili per l'analisi si possono trovare qui [60].

Lo spazio dell'input è un sottoinsieme nel quale rientrano tutti i segnali audio che possono essere ascoltati da uno stetoscopio elettronico posizionato sul petto di un soggetto. È possibile individuare alcune classi di suoni nello spazio dell'input in base alle sorgenti:

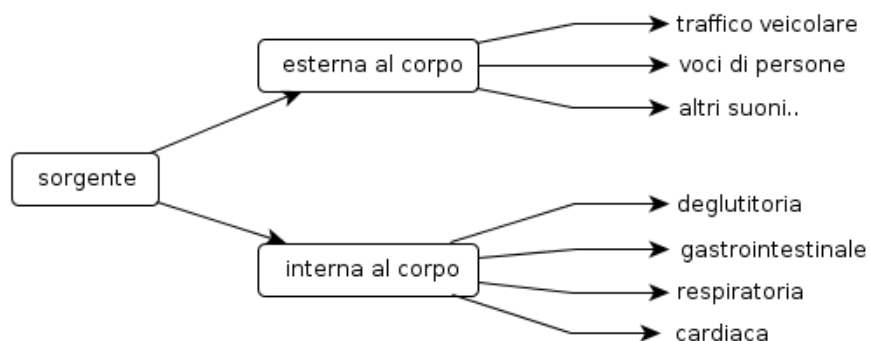


Figura 7.1: Sorgente dei suoni rilevati da uno stetoscopio

## 7.2 Test valutativi

Dopo essersi procurati una vasta gamma di input di suoni respiratori è possibile eseguire alcuni test valutativi per verificare l'effettiva efficienza del sistema, come per esempio:

- Analizzare un file audio lungo, contenente molti cicli respiratori.
- Controllare che un sistema riesca a riconoscere il respiro in una situazione normale sia in una situazione di rumore molto forte.
- Preparare dei file contenente del rumore esterno molto forte e sottoporli al sistema; controllare che questi vengano opportunamente filtrati.
- Dare in input al sistema un file audio contenente solo rumore e controllare che non venga confuso come respiro.
- Aggiungere un suono cardiaco al file e controllare che il sistema riesca a filtrarlo e a non confonderlo con il respiro.
- Simulare una situazione di apnea e vedere se il sistema riesce a riconoscerla in modo efficiente.
- Testare l'efficienza in presenza di white noise.

Se si vuole valutare un algoritmo che localizza nel tempo le fasi respiratorie allora si considera un suono respiratorio come a una sequenza di cicli respiratori e a un ciclo respiratorio come una fase di inspirazione seguita da una fase di espirazione seguita da una pausa. Lo spazio di input è quindi classificabile in sequenze di numeri reali che indicano la posizione temporale dell'inizio di ciascuna fase. L'output dell'algoritmo sarà una sequenza di numeri reali che indicano la localizzazione temporale dell'inizio di ciascuna fase.

Definiamo la differenza tra l'output e il descrittore della classe di cui fa parte l'input come la sequenza dei valori assoluti delle differenze delle singole componenti. La qualità dell'algoritmo si può misurare in termini di qualche proprietà statistica di questa differenza, a esempio la media.

## **Parte III**

### **Conclusioni**



# Conclusioni

In questa tesi si è affrontato il problema del riconoscimento di apnea notturna grazie all'analisi del suono audio della respirazione, registrato attraverso uno stetoscopio elettronico.

Durante la ricerca nella letteratura si è appreso che non sono finora presenti sistemi in grado di soddisfare le nostre richieste. Lo scopo di questa tesi infatti è quello di fornire le premesse necessarie per poter realizzare un sistema in grado di riconoscere in real time le apnee notturne del paziente e dare un allarme in caso della loro presenza.

Una prima analisi dei suoni respiratori registrati ci permette di vederli come una sommatoria del suono del cuore, dei suoni polmonari e dei rumori di sottofondo. Grazie a questa analisi si può subito intuire che prima di procedere con un algoritmo di riconoscimento del respiro bisogna concentrarsi sull'eliminazione dei suoni che possono interferire con la nostra analisi : i rumori di sottofondo e il suono del cuore. Questa prima fase di filtraggio risulta molto importante in quanto se eseguita con poca attenzione condizionerà in negativo il restante procedimento portando per esempio a dei falsi positivi: il sistema non riuscirà a riconoscere la mancanza di respiro in quanto considererà come suoni respiratori altri suoni appartenenti o al battito cardiaco o ad eventuali rumori di sottofondo.

Dopo un'analisi della letteratura si può notare che un possibile approccio per la risoluzione del problema è trattare il suono respiratorio come se fosse musica in quanto anch'esso ha un certo ritmo e quindi risulta ottimale l'uso della tecnica di beat detection per individuarne il ritmo. Grazie all'analisi delle linee guida vista nei capitoli precedenti il collega Federico Viscomi è stato in grado di implementare un sistema per la risoluzione di questo problema ed è così composto:

La tecnica che si consiglia e che è stata usata è un'analisi in real time. Vengono analizzati di volta in volta dei segmenti di suono di una dimensione di quattro secondi. Viene ridotta la frequenza di campionamento per poter migliorare l'efficienza delle fasi successive e vengono applicati dei filtri sia per il denoising sia per eliminare i vari rumori esterni e interni al paziente. In particolare vengono applicati dei filtri passa banda, viene calcolato il valore assoluto e viene applicato un filtro mediano. Dopo le operazioni di

filtraggio viene applicato un algoritmo di beat detection confrontando i campioni con una finestra di analisi di 10 ms. Dato che un beat rappresenta un picco di energia del suono, un metodo semplice per il beat detection è quindi la ricerca dei picchi di energia. Si sceglie di confrontare l'energia istantanea di una finestra del segnale con la media dell'energia di un blocco al più di 4 secondi. Di conseguenza si potrà notare un beat e quindi un respiro quando l'energia istantanea sarà superiore dell'energia media restituendo così un booleano che rappresenta la presenza di un respiro o meno. Dopo questa fase viene usato un algoritmo di breathing pattern recognition che consiste nel clustering dei dati della fase precedente. Ci sono solo due tipi di cluster: presenza di respiro e assenza di respiro, con una soglia d'errore pari al 10%. A questo punto si riescono a distinguere le fasi di respirazione (inspirazione ed espirazione) dalle fasi di pausa e si riesce quindi a riconoscere quando le pause sono troppo lunghe il che significa che si è in una fase di apnea.

La realizzazione di un sistema di riconoscimento di apnea notturna in real time tramite uno stetoscopio elettronico è da considerarsi un ottimo strumento per i pazienti in quanto rappresenta un'alternativa ai metodi esistenti al giorno d'oggi. Attualmente per risolvere il problema dell'apnea viene spesso usata una macchina chiamata PAP [38] dotata di una mascherina nasale o facciale che attraverso un tubo collegato alla macchina eroga aria a una determinata pressione in base alle esigenze del paziente. La PAP resta il trattamento più efficace, soprattutto per le apnee gravi, ma molti pazienti hanno una scarsa aderenza al trattamento perché non usano la maschera per più di quattro ore a notte [56]. Il riconoscimento dell'apnea tramite lo stetoscopio e la PAP hanno due metodologie e due scopi diversi in quanto la PAP previene le apnee, riuscendo anche a eliminarle durante il suo utilizzo mentre il sistema che sfrutta lo stetoscopio elettronico emette un segnale di allarme in caso di mancanza del respiro per svegliare il paziente ed evitare gravi conseguenze. In conclusione, in determinati casi un sistema di riconoscimento dell'apnea tramite uno stetoscopio elettronico può rappresentare una valida alternativa, in quanto è sufficiente attaccare la ventosa dello stetoscopio al paziente, escludendo quindi tutti i fastidi dovuti dalla mascherina della PAP e dalla sua costante erogazione di aria.

## 7.3 Sviluppi futuri

La ricerca nell'ambito della classificazione dell'apnea notturna può procedere in diverse direzioni. Considerando il sistema descritto sopra si può pensare a una sua ottimizzazione come per esempio:

- Migliorare il riconoscimento delle fasi respiratorie e delle apnee per migliorarne la precisione considerando i possibili diversi schemi di respirazione anormale visti nel

capitolo 1.7.

- Aggiungere dei meccanismi di tolleranza al rumore esterno. Ad esempio usare due microfoni uno che registra il rumore ambientale e uno che registra i suoni respiratori e cercare di capire come il rumore esterno modifica il suono respiratorio. Questo problema si può affrontare sia attraverso l'uso di modelli acustici del corpo umano che attraverso esperimenti.
- Portare l'applicazione su un dispositivo mobile permettendo così un uso domestico ancora più semplice.
- Infine, escludendo la metodologia risolutiva appena analizzata, si può considerare un sistema basato su reti neurali per un miglior riconoscimento delle fasi respiratorie e delle apnee, a discapito della portabilità su sistemi mobili.





# Capitolo 8

## Bibliografia

- [1] Laiali Almazaydeh, Miad Faezipour, Khaled Elleithy *A Neural Network System for Detection of Obstructive Sleep Apnea Through SpO2 Signal Features* Department of Computer Science and Engineering University of Bridgeport Bridgeport, CT 06604, USA
- [2] Gina Ann Yi. *A Software Toolkit for Acoustic Respiratory Analysis*. Massachussetts institute of technology.
- [3] Arzt M, Young T, Finn L, Skatrud JB, Bradley TD. *Association of sleep disordered breathing and the occurrence of stroke*. Toronto General Hospital, University Health Network, 9N 943, 200 Elizabeth Street, Toronto, ON M5G 2C4, Canada.
- [4] Juan Pablo Bello, Laurent Daudet, Samer Abdallah, Chris Duxbury, Mike Davies, and Mark B. Sandler. *A Tutorial on Onset Detection in Music Signals*.
- [5] Antonietta Bisulli. *Sindrome delle apnee ostruttive nel sonno(osas): effetti cognitivi del trattamento con pressione continua positiva(CPAP)*. Dottorato di ricerca. Università di Bologna.
- [6] Albert Boggess, Francis J. Narcowich. *A first course in wavelets with Fourier analysis*.
- [7] Joo S. Chuah , Zahra K. Moussavi. *Automated respiratory phase detection by acoustical means*. Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Manitoba, R3T 2N2, Canada.
- [8] G. Charbonneau, E. Ademovic, B.M.G. Cheetham, L.P. Malmberg, J. Vanderschoot, A.R.A. Sovijärvi. *Basic techniques for respiratory sound analysis*.

- [9] S.K. Chowdhury, A.K.Majumder. *Digital spectrum analysis of respiratory sounds*. IEEE Transactions on Biomedical Engineering. Date of Publication: Nov. 1981. Volume: BME-28, Issue: 11 Page(s):784-788
- [10] Rami Cohen. *Signal Denoising Using Wavelets*. Department of Electrical Engineering Technion, Israel Institute of Technology.
- [11] Carlo Drioli, Nicola Orio, *Elementi di Acustica e Psicoacustica*
- [12] Apoor S. Gami, M.D., Daniel E. Howard, B.S., Eric J. Olson, M.D., and Virend K. Somers, M.D., Ph.D. *Day Night Pattern of Sudden Death in Obstructive Sleep Apnea*. New England Journal of Medicine 2005; 352:1206-1214.
- [13] N. Gavriely, Y. Palti, G. Alroy. *Spectral characteristics of normal breath sounds*. Journal of applied physiology 1981 Feb;50(2):307-14.
- [14] Ronald B. George, Richard A. Matthay, Michael A. Matthay, Richard W. Light. *Chest Medicine: Essentials of Pulmonary and Critical Care Medicine. Edition 5*. Lippincott Williams and Wilkins.
- [15] C. Richard Johnson, Jr, William A. Sethares, Andrew G. Klein. *Software Receiver Design: Build Your Own Digital Communication System in Five Easy Steps*. Cambridge University Press. p. 417.
- [16] C. Kandaswamy. *Neural classification of lung sounds using wavelet coefficients*. Computer in biology and medicine, Volume 34, Issue 6, September 2004, Pages 523 - 537.
- [17] H. Klar Yaggi, M.D., M.P.H., John Concato, M.D., M.P.H., Walter N. Kernan, M.D., Judith H. Lichtman, Ph.D., M.P.H., Lawrence M. Brass, M.D., and Vahid Mohsenin, M.D. *Obstructive Sleep Apnea as a Risk Factor for Stroke and Death*. New England Journal of Medicine 2005; 353:2034-2041N.
- [18] Zbigniew Korona, Mieczyslaw M. Kokar. *Lung sound recognition using model theory based feature selection and fusion*. Applied Signal Processing (October 1998), 5(3), pg. 152-169.
- [19] Kompis, M. *Distribution of inspiratory and expiratory respiratory sound intensity on the surface of the human thorax*. Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE.

- [20] Saiful Huq, Azadeh Yadollahi, Zahra Moussavi. *Breath Analysis of Respiratory Flow using Tracheal Sounds*. Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada.
- [21] *Acoustical respiratory signal analysis and phase detection*. S. Le Cam, Ch. Collet, F. Salzenstein. Université Strasbourg 1.
- [22] Yee Leng Yap, Zahra Moussavi, *Respiratory onset detection using variance fractal dimension*, Department of Electrical Engineering, University of Manitoba.
- [23] Paolo Marcellini, Carlo Sbordone. *Analisi matematica uno*.
- [24] R.J.E. Merry. *Wavelet Theory and Applications. A literature study*.
- [25] Roberto Munoz, Joaquín Duran-Cantolla, Eduardo Martínez-Vila, Jaime Gallego, Ramón Rubio, Felipe Aizpuru and Germán De La Torre. *Severe Sleep Apnea and Risk of Ischemic Stroke in the Elderly*.
- [26] Z. Moussavi, M. Leopando, H. Pasterkamp, G. Rempel. *Computerized acoustical respiratory phase detection without airflow measurement*. Medical and Biological Engineering and Computing 2000, Volume 38, Issue 2, pp 198-203.
- [27] NHLBI: Health Information for the Public. *Sleep Apnea: What Is Sleep Apnea?*. U.S. Department of Health and Human Services.
- [28] Cristina Occhipinti *Analisi dei segnali audio mediante funzioni wavelet*
- [29] Hans Pasterkamp, Steve S. Kraman, George R. Wodicka. *Respiratory sounds: advances beyond the stethoscope*.
- [30] Jussi Pekonen. *Onset detection methods for musical sounds*. Helsinki university of technology.
- [31] Wojciech Pieczynski. *Multisensor triplet Markov chains and theory of evidence*.
- [32] Susan Redline<sup>1</sup>, Gayane Yenokyan<sup>2</sup>, Daniel J. Gottlieb<sup>3,4</sup>, Eyal Shahar<sup>5</sup>, George T. O'Connor<sup>3</sup>, Helaine E. Resnick<sup>6,7</sup>, Marie Diener-West<sup>2</sup>, Mark H. Sanders<sup>8</sup>, Philip A. Wolf<sup>3</sup>, Estella M. Geraghty<sup>9</sup>, Tauqeer Ali<sup>9</sup>, Michael Lebowitz<sup>11</sup>, and Naresh M. Punjabi. *Obstructive Sleep Apnea Hypopnea and Incident Stroke*.
- [33] Rosqvist T, Paajanen E, Kallio K, Rajala HM, Katila T, Piirilä P, Malmberg P, Sovijärvi A. *Toolkit for lung sound analysis*. Department of Technical Physics, Helsinki University of Technology, Finland.

- [34] Sergios Theodoridis, Konstantinos Koutroumbas. *Pattern recognition*. Elsevier Academic Press.
- [35] M J Tobin; T S Chadha; G Jenouri; S J Birch; H B Gazeroglu; M A Sackner. *Breathing patterns. 1. Normal subjects..* Chest Journal. 1983;84(2):202-205.
- [36] Yildirim I, Ansari R, Moussavi Z. *Automated respiratory phase and onset detection using only chest sound signal*. University of Illinois at Chicago, IL 60607, USA.

## 8.1 Sitografia

- [37] <http://atlantemedicina.wordpress.com/2008/11/04/auscultazione-del-torace/>  
data di accesso: 28/01/2013.
- [38] <http://it.wikipedia.org/wiki/C-PAP>  
data di accesso: 26/2/2013.
- [39] <http://it.wikipedia.org/wiki/Clustering>  
data di accesso: 6/2/2013.
- [40] [http://commons.wikimedia.org/wiki/File:Wavelet\\_-\\_Morlet.png](http://commons.wikimedia.org/wiki/File:Wavelet_-_Morlet.png)  
data di accesso: 22/2/2013.
- [41] Analisi dei segnali campionati.  
[www.diee.unica.it/misure/Dispense/Misure\\_Elettroniche\\_dm270/Analisi\\_di\\_segnali\\_campionati.pdf](http://www.diee.unica.it/misure/Dispense/Misure_Elettroniche_dm270/Analisi_di_segnali_campionati.pdf)  
data di accesso: 04/02/2013.
- [42] <http://www.elemania.altervista.org/adda/quant/quant1.html>  
data di accesso: 1/3/2013.
- [43] <http://editoria.wiki-site.com/index.php/DIGITALIZZAZIONE>  
data di accesso: 04/02/2013.
- [44] <http://www.griaf.unipa.it/promonda/pagine/composizione%20e%20scomposizione.htm>  
data di accesso: 18/2/2013.
- [45] [http://it.wikipedia.org/wiki/Filtro\\_mediano](http://it.wikipedia.org/wiki/Filtro_mediano)  
data di accesso 20/1/2013.

- [46] [http://fisicaondemusica.unimore.it/Teorema\\_di\\_Fourier.html](http://fisicaondemusica.unimore.it/Teorema_di_Fourier.html)  
data di accesso: 04/02/2013.
- [47] <http://it.wikipedia.org/wiki/Fourier-Transformation>  
data di accesso 20/1/2013.
- [48] <http://it.wikipedia.org/wiki/Invasività>  
data di accesso: 28/01/2013.
- [49] <http://www.larapedia.com/apparato-respiratorio/apparato-respiratorio-umano.html>  
data di accesso 29/1/2013.
- [50] Corso di elaborazione audio.  
<http://www.mediasystemnet.it/videocorsi.html>  
data di accesso: 04/02/2013.
- [51] <http://www.meddean.luc.edu/lumen/MedEd/medicine/pulmonar/physio/pf11.htm>  
data di accesso: 01/02/2013.
- [52] <http://www.meteorologia.it/wavelet.htm>  
data di accesso: 22/2/2013.
- [53] <http://www.nlm.nih.gov/medlineplus/ency/article/003932.htm>  
data di accesso: 04/02/2013.
- [54] <http://otolabdidattica.wordpress.com/2011/02/23/glossario-campionamento/>  
data di accesso 20/2/2013.
- [55] <http://pacs.unica.it/biblio/fisiopatologia/fisiopatologia1.pdf>  
data di accesso: 28/01/2013.
- [56] <http://www.pedettaortodonzia.it/russamento-apnea-notturna.asp>  
data di accesso 26/2/2013.
- [57] [http://it.wikipedia.org/wiki/Processo\\_markoviano](http://it.wikipedia.org/wiki/Processo_markoviano)  
data di accesso 22/2/2013.
- [58] [http://en.wikipedia.org/wiki/File:Signal\\_envelopes.png](http://en.wikipedia.org/wiki/File:Signal_envelopes.png)  
data di accesso: 04/02/2013.

- [59] *Gli stetoscopi elettronici* Giorgio Carlo Monti, Massimo Fusello  
[http://www.simg.it/Documenti/Rivista/2002/08-10\\_2002/12.pdf](http://www.simg.it/Documenti/Rivista/2002/08-10_2002/12.pdf)  
data di accesso: 04/02/2013.
- [60] Sound repositories:
- [http://solutions.3mitalia.it/wps/portal/3M/it\\_IT/Littmann/stethoscope/education/heart-lung-sounds/](http://solutions.3mitalia.it/wps/portal/3M/it_IT/Littmann/stethoscope/education/heart-lung-sounds/)
  - <http://tracheostomy.com/resources/videos/index.htm>
  - <http://faemse.org/downloads.html>
  - <http://www.meddean.luc.edu/lumen/MedEd/medicine/pulmonar/pd/auditory.htm>
- [61] [http://it.wikipedia.org/wiki/Spirometro#Spirometro\\_con\\_pneumotacografo](http://it.wikipedia.org/wiki/Spirometro#Spirometro_con_pneumotacografo)  
data di accesso: 29/01/2013.
- [62] Francesco Fanelli  
*Dall'analogico al digitale: Evoluzione e storia delle apparecchiature elettroniche per musica*  
<http://www.suonoelettronico.com>  
data di accesso 2/1/2013.
- [63] <http://web.tiscalinet.it/santacecilia/sme/audiodigitale/audiodigitale.html>  
data di accesso 22/2/2013.
- [64] <http://www.treccani.it/vocabolario/pneumotacografo/>  
data di accesso: 29/01/2013.